Universitat Autònoma de Barcelona

Essays on Economics of Inequality: Early Childhood Development and Ethnic Favoritism

Augustin Tapsoba

Universitat Autònoma de Barcelona

Doctoral Thesis

Essays on Economics of Inequality: Early Childhood Development and Ethnic Favoritism

Author: Augustin Tapsoba

Advisors: Prof. Joan Llull Cabrer Prof. Hannes Felix Mueller

A dissertation submitted to the *Departament d'Economia i d'Història Econòmica* in partial fulfillment of the requirements for the degree of Doctor of Philosophy by the *International Doctorate in Economic Analysis* (IDEA).

June 4, 2019

Acknowledgements

I am deeply indebted to my supervisors Joan Llull and Hannes Mueller. They have provided me with invaluable guidance and support during my thesis. Beyond their amazing roles as supervisors, they believed in me more than anyone ever and gave me the means to overcome every bump during this incredible journey. They have been nothing but patient, encouraging and inspiring throughout all these years and no words can truly express my gratitude to them.

I would like to also express my gratitude to all the faculty members and administrative staff of IDEA program for their help and support. The open, friendly and engaging top level research environment they have created has been very helpful to me. I am specially thankful for the countless comments and feedback that I have received all these years. I am especially thankful to the directors of IDEA program during my PhD years Suzanna Esteban, Lucas Gambetti, Francesc Obiols, Inés Macho-Stadler, David Perez for all the academic, financial and logistic support they have secured to put me in the best conditions to write this thesis. I am also thankful to Àngels López Garcia, Mercè Vicente, Marta Molina and Bruno Guallard for their wonderful administrative support at IDEA and Barcelona Graduate School of Economics (GSE).

I gratefully acknowledge financial support from the Spanish Ministry of Economy and Competitiveness through an FPI Grant at Barcelona GSE. I am also thankful to the Suntory and Toyota International Centre for Economics and Related Disciplines (STICERD) at the London School of Economics and Political Science (LSE) for hosting me as a visiting PhD student.

Thank you to all the colleagues and friends from all the cohorts of PhD students that have overlapped with my stay at IDEA. Our social interactions, shared experiences, mutual feedback and support have been very useful to me. I would like to specially thank Bruno Conte Leite and Lavinia Piemontese for co-authoring the second chapter of this thesis with me.

Finally, I would like to express my heartfelt gratitude to my family without whom none of this would have been possible. My life-partner Alena Damkova has been my rock during this journey. Her love, patience, encouragement, feedback and unwavering trust in me coming through with this thesis have motivated me each and every day specially during the toughest ones. I would also like to thank my parents Raphaël and Pauline for their continued support throughout my whole life, for believing in me and for respecting my choice of going into the PhD adventure. To my younger siblings Nadège, Jeanette and Christophe for their loving presence in my life.

Contents

Abstract

In this dissertation, I investigate the causes and consequences of economic inequality along two dimensions: early childhood development for inequality among individuals and ethnic favoritism for inequality among social groups. In three separate essays, I specifically study the consequences on child development of shocks experienced early in life on one hand, and the prevalence and determinants of ethnic favoritism around the world on the other.

Chapter 1 studies the impact of conflict risk on child health and it shows that insecurity in conflict-prone areas can affect child health even in absence of immediate violence. Beyond the damage caused to direct victims of violence, behavioral responses to insecurity can lead to major health setbacks for young children. The fear of exposure to conflict events often triggers such responses even before/without any manifestation of violence in a given area. This generates a treatment status (exposure to conflict risk) that goes beyond violence incidence. In this chapter, I investigate the impact of conflict on child health using a new metric that captures perceived insecurity at the local level through a statistical model of violence. In this model, violence is a space-time process with an unknown underlying distribution that drives the expectations of agents on the ground. Each observed event is interpreted as a random realization of this process, and its underlying distribution is estimated using adaptive kernel density estimation methods. The new measure of violence risk is then used to evaluate the impact of conflict on child health using data from Ivory Coast and Uganda. The empirical evidence suggests that conflict is a local public bad, with cohorts of children exposed to high risk of violence equally suffering major health setbacks even when the risk does not materialize in violent events around them.

Chapter 2 of the thesis is joint work with Bruno Conte Leite and Lavinia Piemontese.

It investigates the impact on child health of the 2004 locust plague invasion in the Sahel region of Africa. It argues that locust invasions in these agricultural economies generate first a speculative/anticipatory price effect during the plague itself followed by a local crop failure effect that could constitute a food supply shock for local markets and an income shock for farmers. Using a Difference-in-Differences identification strategy, we show that children exposed in utero to the adverse effects of locust plagues suffer major health setback. Exposed children have, on average, a Z-score 0.25 points and 0.48 points lower than non-exposed children in Mali and Senegal respectively. Children exposed to the speculative price effect suffer as much as those exposed to the crop failure effect.

Finally, chapter 3 of this thesis is joint work with Hannes Mueller and studies the prevalence and determinants of ethnic favoritism around the world. It uses a dataset which codes executive power for 564 ethnic groups in 130 countries on a seven-point scale to show that ethnic groups that gain political power benefit economically. This effect holds for groups that enter government, the extensive margin, and for groups that concentrate more power onto themselves, the intensive margin. Both these effects disappear in the presence of strong political constraints on executive power. Institutional constraints are even effective in preventing favoritism when groups concentrate all power in the executive onto themselves.

Chapter 1

The Cost of Fear: Impact of Violence Risk on Child Health During Conflict

1.1 Introduction

Violent conflicts are serious humanitarian and economic threats to many developing countries.[1] Their impact on child health can substantially increase the overall cost of conflict and heavily affect the timing and nature of the recovery process. The early years of life are indeed crucial in influencing a range of health and social outcomes across the life-course. In particular, health in the fetal period, infancy, and early childhood determines most of our long-term health, education and labor market outcomes (Black et al., 2007; Behrman and Rosenzweig, 2004; Maccini and Yang, 2009; Maluccio et al., 2009; Stein et al., 1975). Conflict-induced shocks experienced at this stage of life can, therefore, have persistent detrimental effects that we need to understand in order to devise the proper responses to protect the next generations in conflict-prone environments.

Beyond the damage caused to direct victims of violence, conflict can lead to major health setbacks for young children through the consequences of insecurity. Economic

[1] The Internal Displacement Monitoring Centre estimated, for instance, that 2.7 million people were newly displaced in Africa during the first six months of 2017, and three-quarters of those were allegedly due to conflict.

agents react to violence risk with displacements or changes in consumption/production behavior that can lead to disruptions in supply and demand of goods and services. They sometimes engage in such risk coping behavior before/without any manifestation of violence in a given area based on their expectations. In order to properly evaluate the consequences of conflict, it is, therefore, crucial to capture the risk perceived by agents on the ground beyond the incidence of violent events.

This paper investigates the impact of conflict on child health using a new metric that captures the perceived insecurity at the local level through a statistical model of violence. It is built on the intuition that economic agents react to the risk of being exposed to violence, and this generates a treatment status (exposure to conflict risk) which is unobserved but crucial for the short and long-term effects of conflict. In order to build a new metric that captures this treatment beyond the incidence of conflict events, violence is modeled as a space-time process with two characteristics. First, the occurrence of an event increases the likelihood that another event occurs in the same area. Second, violence can spread through the local environment via contagion.[2] This process has an unknown distribution that, by assumption, drives the expectations of agents on the ground. The density of this distribution is backed out of the observed pattern of violent events across space and over time.

In order to estimate this density, a locally adaptive kernel method is implemented on the observed sample of violent events.[3] The basic principle behind this approach is that each observed event has its own contribution to the overall density at a given location. This contribution is given by a trivariate Gaussian kernel function. It is highest at the exact location of the event and fades out as we move away from it in space or time. The dispersion of the kernel mass is controlled by a matrix of smoothing parameters, with flatter kernels for events in low intensity areas. The density at a given location is obtained

[2]An illustration of the first case is when battles between organized actors trigger waves of retaliatory or follow-up violence in a given area. For the second case, troops repeatedly attack clusters of nearby targets. This may happen because local vulnerabilities are well-known to them. Moreover, armed combatants can easily migrate from one area to another and violence in a given city can disrupt regional economic stability.

[3]The observed events are interpreted as a sample of random realizations of the underlying violence process. The density estimation approach tries to estimate its unknown density based on the location of observed events in space and time. It is conceptually the opposite of drawing a sample from a given distribution.

by averaging all the contributions.

This new measure of violence risk is then used to evaluate the impact of conflict on child health using data from Ivory Coast and Uganda.[4] The probability of exposure to violence in utero or during the first year of life across space for different cohorts of children is computed in order to estimate the impact of violence risk on infant mortality. The identification strategy relies on the spatial and temporal variation in the exposure of different birth cohorts to violence in Ivory Coast and Uganda, in a Difference-in-Differences setting. The distribution of events in space and time generates space-time windows with high risk of violence but no violent events and windows with isolated events but low risk of violence.

My main finding is that violence risk significantly increases infant mortality even when the risk does not materialize in actual violence: a standard deviation increase in the probability of being exposed to violence between conception and the first anniversary increases infant mortality by 1 and 0.8 percentage points in Ivory Coast and Uganda, respectively. The estimated effect is quantitatively large. Children in the top risk quartile experience an increase in infant mortality of 6 and 3 percentage points in Ivory Coast and Uganda, respectively, which represents more than half of the average infant mortality rates in both countries. Interestingly, this effect is similar in magnitude and significance level when the analysis is restricted to the cohorts of children that had no violent incident happening around their area of birth. Relying on the incidence of conflict events to define exposure to the adverse effects of conflict leads to an underestimation of the share of children that are treated and the magnitude of the treatment effect. The combined effect suggests that the standard metrics are unable to account for up to 70% of all the conflict-related child deaths. The empirical evidence in Uganda also suggests that quality/availability of health care services, malnutrition and maternal stress are important transmission channels for this effect.

The baseline regressions control for household head, mother and child characteristics, weather shocks, enumeration area, and birth cohort fixed effects. The estimates are robust

[4]Ivory Coast has experienced a relatively low-intensity but highly disruptive conflict between 2002 and 2011. Uganda experienced an exogenous outburst of violence between 2002 and 2005 that put an end to a long-lasting ethnic conflict.

to many different specifications and sample restrictions, including a very demanding one that controls for mother fixed effects. Placebo tests confirm that results are not driven by preexisting trend differences between high and low risk areas.

The evidence documented above suggests that conflict is a local public bad and being exposed to high risk of violence leads to major setbacks in child health, even when the risk does not materialize in actual violence.[5] This is in line with the idea that economic agents engage in detrimental ex-ante coping strategies that lead to disruptions in the supply and demand of goods and services at the local level. From the policy perspective, this suggests that we should rethink both the scale and the target of interventions that aim at mitigating the consequences of conflict.

This paper contributes to three main strands of the economics literature. First, it contributes to the previously discussed literature on the importance of early life conditions (Lavy et al., 2016; Maccini and Yang, 2009; Maluccio et al., 2009; Black et al., 2007; Behrman and Rosenzweig, 2004; Stein et al., 1975). A substantial part of this literature has specifically looked at the effects of conflict on child health. It has established that exposure to conflict leads to worse birth outcomes (Mansour and Rees, 2012; Camacho, 2008), increases infant and child mortality (Dagnelie et al., 2018; Valente, 2015), and decreases height-for-age z-scores (Akresh et al., 2016; Minoiu and Shemyakina, 2014; Akresh et al., 2012b; Bundervoet et al., 2009). All these papers rely on the incidence of violent events to proxy exposure to the adverse effects of conflict.[6] They were, therefore, unable to account for any of the cost induced by behavioral responses to risk in absence of immediate violence. Counting the number of events that happen in a given space-time window as a measure of conflict risk implies that violence risk is high only for windows in which we already observe violent events. In other words, violence is perceived to be likely to happen only where it has already happened. The cohorts that are not exposed to any event in their city or village in a given time period are therefore considered to be

[5]Part of this result could also be coming from the fact that the new metric is more robust to measurement errors in the conflict event data. The estimation of the underlying density of the data generating process does not require the universe of events but just a representative sample and space-time windows with isolated or misreported events will still have a low conflict risk.

[6]The standard approach in the literature is to count the number of events or fatalities within a given space-time window. Papers that link conflict to other outcomes also use this approach (Ouili, 2017; Shemyakina, 2015; Leon, 2012; Akresh et al., 2012a; De Walque, 2005).

equally non-treated irrespective of the actual risk perceived by the agents on the ground. The new approach proposed in this paper is more general and fixes these limitations by trying to capture violence risk beyond the incidence of conflict events. It uses the pattern of these events across space and over time to estimate the underlying distribution of the process that generated them. This is also the first paper to provide evidence suggesting that children exposed to high risk of violence suffer major health setbacks even when this risk does not materialize in actual violence around them.

Second, this paper also contributes to a small but growing literature that tries to evaluate the consequences of insecurity beyond the incidence of violent events in conflict and crime literature. In the conflict literature, Rockmore (2017) shows that subjective risk (perceptions of survey respondents on difficulties to cultivate land due to insecurity) has a higher impact than violence incidence (actual attacks against a community) on household consumption in Northen Uganda. In fact, he finds that half of welfare losses caused by conflict are related to risk and not to direct exposure to violence. Arias et al. (2014) investigate the changes in household production behavior and show that conflict affects land use and agricultural investment beyond violence incidence. Using data from Columbia, they show that during the initial years of presence of armed groups, farmers cut back production of perennial crops and pasture in favor of seasonal crops as a coping strategy to insecurity. My paper is built on similar intuitions but the methodological approach used to capture risk is very different. I propose a framework in which violence risk can be derived from the observed location of violent events in space and time. Moreover, this paper is the first one within this literature to look at child health as an outcome. The existing papers have only looked at changes in household consumption or production behavior (Rockmore, 2017; Arias et al., 2014; Bozzoli and Brück, 2009; Deininger, 2003).

The idea that risk perception goes beyond incidence has also been addressed in the crime literature. The fear of crime has been shown to affect housing prices even in absence of any criminal event. Pope (2008) uses registry data that tracks sex offenders in Hillsborough County, Florida, and shows that nearby housing prices fall (rebound immediately) after a sex offender moves into (out of) a neighborhood irrespective of the seriousness of the threat that they pose. Fear of crime has also been measured with self-reported data on perceived insecurity (Buonanno et al., 2013) and urban property crime rates (Gibbons, 2004) to show its impact on housing prices. From the methodological perspective, kernel density estimation has been widely adopted in crime literature for hotspot mapping and crime prediction (Hu et al., 2018; Gerber, 2014; Chainey et al., 2008). The idea of using it to capture perceived risk of crime on the ground is however, to the best of my knowledge, new to this literature. One can study the consequences of behavioral responses to a specific type of criminal events beyond their incidence following the methodology proposed in this paper.

The third strand of literature related to my work looks at the role of expectations in conflict-prone environments. At aggregate level, Zussman et al. (2008) and Willard et al. (1996) show, for instance, that asset prices during conflict react to important conflict events like battles or ceasefire agreements.[7] Besley and Mueller (2012) study the effect of violence on house prices in Northern Ireland. They show that the peace process and its corresponding decline in violence led to an increase of house prices in the most affected regions compared to other regions. Their paper offers a way to understand the heterogeneity of these changes across regions which directly links to the role of expectations at local level. The current paper contributes to this literature by showing that conflict can also affect child health and long term economic development at local level through expectations.

The remainder of the paper is organized as follows. The next section gives some background to the Ivorian and Ugandan conflicts and describes the data. Section 1.3 presents a methodological approach for measuring violence risk across space and time. Section 1.4 presents the identification strategy for estimating the impact of conflict on infant mortality, followed by the results and some robustness checks in Section 1.5. Section 1.6 concludes.

[7] See Mueller et al. (2017) for a full discussion on the role of expectations in conflict affected countries.

1.2 Historical Background and Data

1.2.1 Data Sources

To locate violence in space and time, I use information from conflict event datasets (ACLED and GED) on the exact dates and locations of violent incidents during the conflict, including armed battles, and violence against civilians.[8]

To assess the impact of exposure to violence risk on child health, I use Demographic and Health Survey (DHS) data collected in 2011-2012 in Ivory and 2006, 2011 and 2016 in Uganda. DHS surveys are a set of repeated cross-sectional surveys run in most developing countries since late 1980s. It gathers information on demographic topics such as fertility, child mortality, health service utilization, and nutritional status of mothers and young children from a nationally representative sample of households. Women (aged 15 to 49) are interviewed about the birth and survival of almost all the children they gave birth to in the past (up to 20 children), including those who died by the time of the interview. For the main analysis, I look exclusively at the survival of single born female children.[9] I also use information on use of child health services and maternal health during pregnancies that happened at most 5 years prior to a DHS interview to shed some light on the potential mechanisms through which behavioral responses to violence risk can affect child health.

Given the emerging evidence on the links between weather shocks and violence (Harari and La Ferrara, 2017; Burke et al., 2015; Hsiang et al., 2013), I also introduce climatic variables to reduce risk of confounding factors. Climate data comes from Terrestrial air temperature and precipitation: 1900–2014 gridded time series (Matsuura and Willmott, 2015). This dataset is a compilation of updated sources and provides monthly precipitation (and mean temperature) interpolated to a latitude/longitude grid of 0.5 degree by 0.5 degree. Following Kudamatsu et al. (2012), I compute average rainfall and temper-

[8] I exclude riots and protests from the conflict data and I keep all the events recorded with geographical precision at city level or lower. Duplicates (same location and time) are eliminated to have a count of the number of days with at least 1 event for a given location.

[9] Multiple birth babies face significantly higher chances of dying before turning one year old. Male fetuses are biologically more fragile compared to female ones and this leads to a higher selection in the sample of born males in presence of shocks around pregnancy period (Dagnelie et al., 2018; Pongou, 2013).

ature for the last two years before birth and for the first year of life and include them separately in regressions to control for variations in maternal food supply induced by climate shocks.

1.2.2 Conflict in Ivory Coast

Ivory Coast is a previous French colony that enjoyed a prolonged period of economic growth since its independence in 1960 until 1990. Political instability has been sparked by the power struggle following the death of the country's first president, Felix Houphouet-Boigny, in 1993. The Interim President, Henri Konan Bedie, in order to secure a win in the 1995 elections, changed the electoral code to exclude his rival and former Prime Minister, Alassane Dramane Ouattara, from running based on his status of son of migrant. The 1995 election was then boycotted by the other opposition leaders in protest against this discrimination and Bedie was elected with 96% of the votes.

In July 1999, Alassane Ouattara, left his job at the International Monetary Fund and returned to run for the 2000 presidential elections. His plan to challenge Bedie again splits the country along ethnic and religious lines based on their own origins which determined their electoral basis. This political turmoil gave a pretext to a group of army soldiers to intervene and overthrow Bedie by a coup in December 1999. Robert Guei, an army officer was chosen to lead a transition to new elections and restore the order. New elections took place in October 2000, but Alasane Ouattara was still excluded from the electoral process because of not being "Ivorian enough". General Guei proclaimed himself president after announcing he had won the presidential elections, but was forced to flee in the wake of a popular uprising and was replaced by his challenger Laurent Gbagbo. Fighting erupted between President Gbagbo's mainly southern Christian supporters and followers of his main opponent Alasane Ouattara, who were mostly Muslims from the north. Tensions lasted for almost a year before both challengers agreed to work towards reconciliation in March 2001.

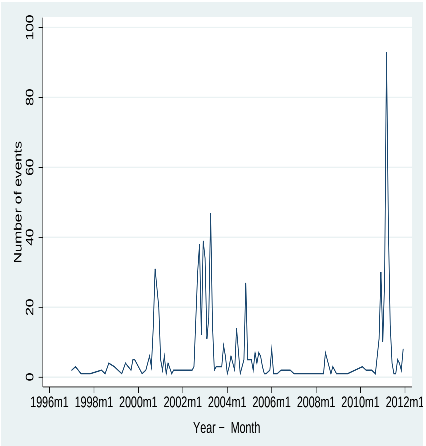
This fragile reconciliation process was disrupted in September 2002 when a mutiny in Abidjan grew into a full-scale rebellion with Ivory Coast Patriotic Movement rebels seizing control of the north. French interposition troops were sent to limit the clashes between the two armies as shown in Figure A2 but intense battles between the two sides took place until March 2003 when the first peace agreement was signed and rebels made their entry into the government. Many other peace talks were held as actors were resuming clashes at one point or another during the implementations of the different peace agreements. This went on until March 2007 when a power sharing deal was signed. Under this deal, Guillaume Soro, leader of the rebel group, was made Prime Minister of Ivory Coast and the new government was put in charge of preparing the elections that would end the crisis and restore a stable constitutional order. The elections after being postponed twice were finally held in December 2010 but led to another crisis. The electoral commission declared Mr Ouattara the winner of presidential election run-off. Mr Gbagbo refused to accept these results and the dispute between the two camps soon escalated into extremely violent clashes until the capture of Gbagbo in April 2011 after the loyalist army has been defeated by the rebels backed by French troops under UN mandate.

Figure 1.1a shows the number of recorded events per month across the whole country. We can see that the recorded events are consistent with the timing described in the preceding paragraphs with some violence peaks after the 2000 elections, during the first months following the start of the first civil war in 2002 and the second civil war following the 2010 elections. Fewer events were recorded during the negotiations period until the comprehensive power sharing treaty was signed in 2007. The spatial distribution of the events is also shown in Figure 1.1b. One can see that conflict incidents are mostly clustered in the central and western parts of the country.

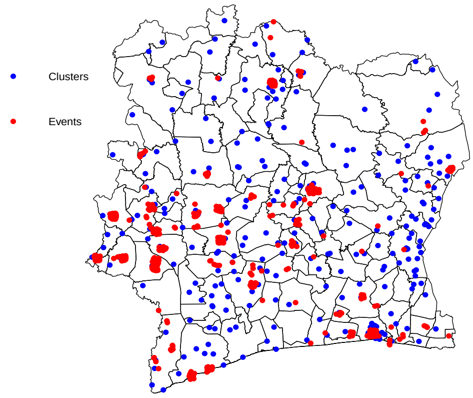
1.2.3 Conflict in Uganga

Uganda has experienced multiple conflicts since its independence in 1962. It has been ruled by the National Resistance Movement (NRM) led by Yoweri Museveni since 1985 whose main constituency is the Bantu-dominated South. His government has faced opposition and armed rebellion in several parts of the country, especially in the North, where the Lord's Resistance Army (LRA) was active until 2006, and close to the border with the Democratic Republic of Congo, where the insurgency led by the Allied Democratic Forces (ADF) was active until 2004.

Figure 1.1: Distribution of Violent Events in Ivory Coast Between 1997 and 2012



(a) Monthly Number of Events

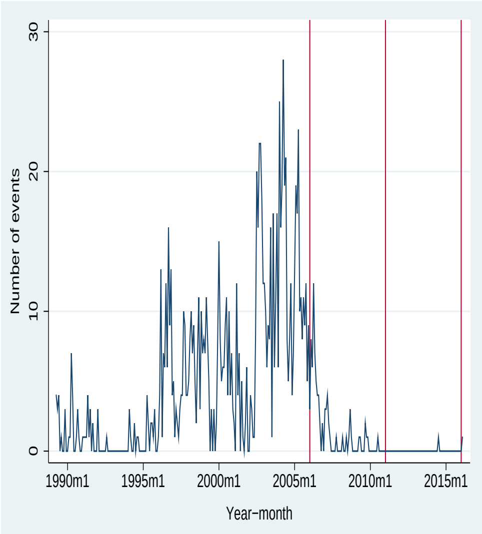


(b) Spatial Distribution of Events and Household Clusters

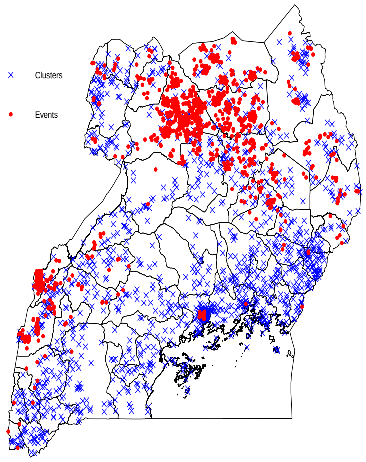
After September 11 terrorist attack, LRA and ADF were declared terrorist organizations by the US Patriotic Act and lost support from their allies that feared retaliations and sanctions. LRA, in particular, lost support from the Sudanese government that was providing them sanctuary and military help. The ruling government in Uganda seized this opportunity to launch a military crackdown on these rebel groups. ADF has been defeated by 2004. Military action against the LRA started in March 2002, when the army launched "Operation Iron Fist" against the rebel bases in South Sudan. The LRA responded by attacking villages and government forces in Northern Uganda. A cease-fire between the LRA and the government of Uganda was signed on September 2006, with the mediation of the autonomous government of South Sudan.

Figure 1.2a shows the monthly number of geo-referenced fighting events conflict between 1988 and 2016 from GED data. Consistent with the narrative above, there was a sharp increase in 2002-05, followed by a decline, and very low levels of violence have been recorded after 2006.

Figure 1.2: Distribution of Violent Events in Uganda Between 1989 and 2016



(a) Monthly Number of Events



(b) Spatial Distribution of Events and Household Clusters

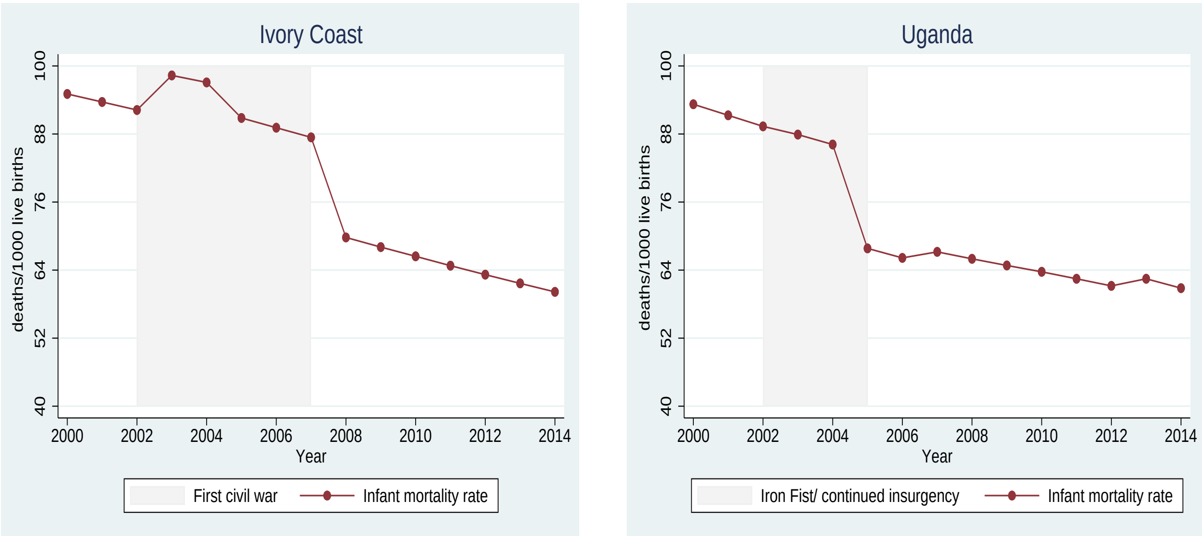
Notes: Red bars in figure (a) are DHS household survey years in Uganda.

1.2.4 Violence and Infant Mortality in Ivory Coast and Uganda

The conflict in both Ivory Coast and Uganda had bad consequences on a wide range of outcomes. Economic activity for instance has been significantly affected. Both countries experienced a prolonged period of negative GDP growth during years of intense violence. Another crucial indicator that has suffered from conflict in both countries is infant mortality rate. This indicator, often used as proxy for the quality of health services in a given country, has gotten worse in Ivory Coast during the first civil war (2002-2007). It has improved substantially after violence ended in Uganda following the government crack down on rebel groups between 2002 and 2005 (see Figure 1.3). Beyond this apparent deterioration of health system during conflict, the aim of this is to provide causal evidence of the impact of behavioral response to violence risk on infant mortality.

The anecdotal evidence suggest that fear of exposure to violence has been one of the key sources of the detrimental impact of conflict on child health. In Ivory Coast for instance, it is estimated that 70 percent of professional health workers and 80 percent of government-paid teachers abandoned their posts in the northern and western parts of the country after the civil war started (UNOCHA, 2004; Sany, 2010). Survey data collected in the western province of Man showed that the three most important conflict-related problems reported by households were health problems (48 percent), lack of food (29 percent), and impaired public services (13 percent) (Fürst et al., 2009). Provision of goods and services has therefore been heavily disrupted in entire regions of the country and for many years due to expectations of economic agents on the ground and beyond actual manifestation of violence. This is what the current paper is trying to capture.

Figure 1.3: Conflict and Infant Mortality



(a) Ivory Coast

(b) Uganda

Source: CIA World Factbook

1.3 Estimation of Violence Risk in Space and Time: Kernel Density Estimation Approach

In order to capture behavioral responses to violence risk, we need a way to quantify risk perceptions at the local level. To do this, I model the observed conflict events as random realizations of an underlying process which is observed by individuals on the ground but has to be backed out of the violence data. In this section, I show how non-parametric density estimation methods can be used to capture variations in violence risk at local level and describe beliefs on the ground.

Non-parametric density estimation methods have been initially used in the literature to assess basic characteristics of an unknown distribution such as skewness, tail behavior, number, location and shape of modes (Silverman, 1986). Nowadays, they play a major role in machine learning, classification and clustering.[10] They are popular methods in crime literature for hotspot mapping and in seismology literature for seismic hazard estimation. They can also be used (like in this paper) as input for more sophisticated analysis. DiNardo et al. (1996) used kernel density estimation method to build counter-factual densities in order to study the effects of institutional and labor market factors on changes in the U.S. distribution of wages in the 80s. Kernel density estimation methods have also been extensively used in poverty analysis to measure poverty from grouped data (mean incomes of a small number of population quantiles) through the estimation of the underlying global income distribution (Sala-i-Martin, 2006; Minoiu and Reddy, 2014; Sala-i-Martin, 2002).

1.3.1 Principle of Kernel Density Estimation (KDE)

A "Simple" Density Estimation

Let's assume violence process X has an unknown probability density function (pdf) f. Density estimation consists of constructing an estimate of f based on a representative sample $\{x_1, ..., x_n\}$ of X.

Let's begin with the simple case of a continuous, univariate random variable X.[11] Let $F(x)$ denote the cumulative distribution function of X. From the definition of the density, we know that

$$f(x) = \frac{d}{dx}F(x) = \lim_{h \to 0} \frac{F(x + \frac{h}{2}) - F(x - \frac{h}{2})}{h},$$

where h is the width of the interval. The simplest estimator would be to count the

[10] For instance, some clustering methods are based on bump hunting, i.e., locating the modes in the density and Bayes classifiers are based on density ratios that can be implemented via density estimation. Applications of density estimation in machine learning and classification are discussed in more depth in the books of Izenman (2008) and Hastie et al. (2009).

[11] To estimate violence risk across space and time, an extension to 3 dimensions (latitude, longitude and time) is shown below.

number of observations around the point x and divide that number by nh. The resulting estimator can be formally written as

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}\mathbb{1}\left(-\frac{1}{2} < \frac{x_i - x}{h} < \frac{1}{2}\right),$$

where $\mathbb{1}(A)$ takes the value 1 if the argument A is true and 0 otherwise.[12] The indicator function can be replaced with the more general notation of a kernel function $k(*)$ and the estimator is now given as

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}k\left(\frac{x_i - x}{h}\right).$$

The previous equation shows that each observed event, x_i, has a given contribution to the density estimate at x. This contribution is only a function of its distance to x. All the points falling within an interval h around x will contribute equally $(1/nh)$ to the density and all the other points will have zero contribution. The density $f(x)$ is just the sum of all the contributions.

Moreover, this "simple" density estimator is equivalent to counting the number of events that occurred during a given time period within a certain distance from x which is similar to the way the current literature measures exposure to conflict. Conflict affected areas are areas with high density while non-affected areas are those with low or zero density. The parameter h is "handpicked" (administrative areas, cells of a given size or rings of a given radius around x), and all the observations falling within a distance h from x contribute equally to the measure of conflict exposure.

However, this approach has the following limitations:

- The distance from location x to a given conflict event is not fully incorporated into this "simple" density estimate since all events falling within a certain distance have equal contributions. One would like events that are closer to x in space or time to have higher contributions to the density, everything else equal.
- Beyond distance, the dispersion of conflict events should also affect the density. More dispersion in observed violent events should lead to a higher risk everywhere

[12] This is the common histogram.

as opposed to very clustered events where the high risk is very localized.

- The trajectory/orientation of the conflict process in space could also matter for the underlying density: At equal distance and dispersion, people living on the trajectory of the violence process will be more threatened as compared to those that are off the trajectory, exactly like the case of cyclones.

The rest of this section will show how the "simple" density estimation approach can be generalized to overcome these limitations.

Standard Kernel Density Estimation

A more sophisticated estimator of the underlying density can be obtained by replacing the indicator function with a smooth kernel function. The Gaussian functional form is a standard choice in the economics literature.[13] In this case, a Gaussian function $k \sim \mathcal{N}(0, h)$ is placed over each data point to measure its contribution to the density for all the locations around. The value of the kernel function is highest at the location of the data point, and diminishes at an exponential rate as we move away from it. The density estimate at x is obtained by summing up all the contributions. Closer events have higher contributions to the density and clustered events generate peaks corresponding to the modes of the underlying density function as shown in Figure 1.4.

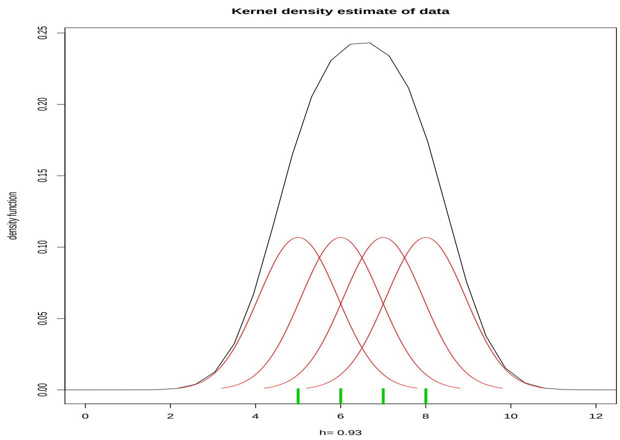
Importance of Smoothing Parameter

The choice of the smoothing parameter h is crucial in kernel density estimation. Small values of h reduce the bias by putting all the mass just around the data point and the density estimate displays spurious variations in the data. When h is too big, each individual kernel becomes flatter and important details in the distribution can be obscured (see Appendix Section A for an illustration).

The choice of the smoothing parameter

[13] The bandwidth is the most crucial choice to make in KDE because it controls the degree of smoothing applied to the data. The kernel form is only responsible for the regularity of the resulting estimate (continuity, differentiability) and Gaussian kernels give estimated density function that has derivatives of all orders (Silverman, 1986).

Figure 1.4: Aggregation of individual kernels in KDE



Notes: Kernel density estimation with 4 data points (in green). The Red Gaussian functions are individual kernels aggregated to get the density estimate in black.

Optimal h is chosen to minimize Mean Integrated Squared Error (MISE)

$$\begin{aligned} MISE(h) &= \int [\hat{f}(x,h) - f(x)]^2 dx \\ &= \int \hat{f}(x,h)^2 dx - 2\int \hat{f}(x,h) f(x) dx + \int f(x)^2 dx, \end{aligned}$$

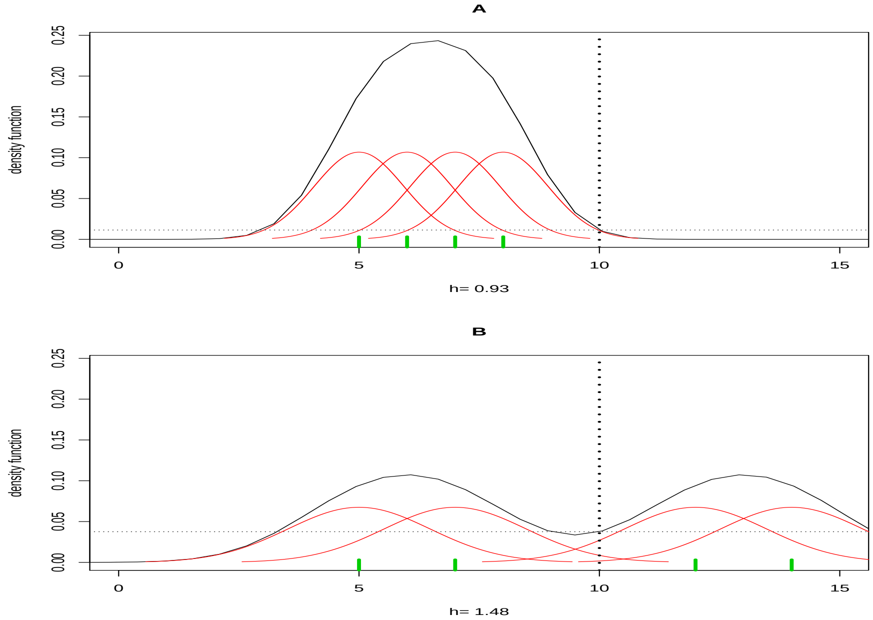
where $\hat{f}(x,h)$ is the kernel density estimate and $f(x)$ is the unknown density.

$MISE(h)$ can be evaluated and minimized without knowing explicitly $f(x)$ by using a bootstrap method (Taylor, 1989) as described in Section B of the appendix.

The dispersion of the observed sample of events is a feature of the underlying density, so minimising the $MISE(h)$ leads to a larger optimal h for samples with more dispersion. The smoothing parameter is indeed an increasing function of the sample variance.[14] This is illustrated in Figure 1.5 where both panels show individual kernels and density estimates from 4 data points at equal distance from a given location $x = 10$. There is less dispersion in the event data in panel A compared to panel B and location x should be more at risk in B than A. The kernel density estimation is able to distinguish these two cases when estimating them as two separate processes.

[14]In the particular case of the underlying pdf being normally distributed, one can show that h_{opt} is actually proportional to sample variance Silverman (1986)

Figure 1.5: Fixed KDE and Sample Variance in Separate Processes



1.3.2 Adaptive KDE

Fixed bandwidth KDE has been shown to perform poorly with densities that exhibit large changes in magnitudes, multi-modalities and long tails (Silverman, 1986). In multi-dimensional cases, these issues are even more exacerbated by the scarcity of the data over much of the effective estimation space (Terrell and Scott, 1992). Allowing for more flexibility in the choice of smoothing parameters is therefore useful in order to estimate the density of an underlying violence process. This can be performed by using adaptive kernel density estimation method.

This method combines features of kernel density estimation and nearest neighbour approach. The idea is to construct a kernel estimate consisting of bumps or kernels placed at observed data points, but allowing the smoothing parameter of the kernels to vary from one point to another. An observation in a low intensity area will therefore have its mass spread out over a wider range than one in a high intensity area (see Terrell and Scott (1992) for more details). The adaptive kernel approach is constructed in a three stage procedure:

- First, find a pilot estimate $\tilde{f}(x)$ using classic kernel density estimation with fixed bandwidth $\tilde{f}(x) = \frac{1}{nh_p}\sum_{i=1}^{n} k\left(\frac{x-x_i}{h_p}\right)$, where $k \sim \mathcal{N}(0, h_p)$ is a Gaussian kernel function.

- Second, define a local bandwidth factor λ_i by $\lambda_i = [\tilde{f}(x_i)/g]^{-\alpha}$, where g is the geometric mean of the $\tilde{f}(x_i)$ and α is a sensitivity parameter such that $0 \leq \alpha \leq 1$.

- Finally, define the adaptive kernel estimate $\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{h\lambda_i}k(\frac{x-x_i}{h\lambda_i})\right)$

The optimal h is chosen to minimize the distance between the estimated density and the unknown density holding local bandwidth factors (λ_i's) constant. This can be done by bootstrap method as in the case of fixed bandwidth.

Allowing the local bandwidth factor to depend on a power of the pilot density gives some flexibility in the design of the method. The larger the power α, the more sensitive the method will be to variations in the pilot density, creating more differences in the smoothing parameters across different parts of the distribution. $\alpha = 0$ is equivalent to fixed bandwidth method while $\alpha = 1$ coincides with the nearest method approach. $\alpha = 1/2$ is the standard choice in the statistics literature (see Terrell and Scott (1992)).

1.3.3 Generalisation to more than One Dimension

The univariate adaptive kernel density estimation can easily be extended to multivariate case. The estimated density is given by

$$\hat{f}(\bar{x}) = \sum_{i=1}^{n} \frac{1}{n \mid H_i \mid} K\Big(H_i^{-1}(\bar{x}_i - \bar{x})\Big),$$

where $\bar{x}_i$ and $\bar{x}$ are d-dimensional vectors, $H_i = \lambda_i H$ is a dxd symmetric matrix of parameters to be estimated and K is a multivariate kernel function.

To estimate conflict risk, we need to consider at least 3 dimensions: latitude, longitude and time. The matrix H can be parametrized as follows:[15]

[15]In the main specification K is a restricted trivariate function in which events that happen beyond 200 Km or after a period t have zero contribution to the density at t. Most of the findings are also robust if these restrictions are ignored.

$$H = \begin{pmatrix} h_s & 0 & 0 \\ 0 & h_s & 0 \\ 0 & 0 & h_t \end{pmatrix}.$$

The integral of the estimated density of the underlying violence process gives an estimate of the probability of occurrence of an event in a given space time window. This is used as risk measure in the rest of the paper.

1.4 Identification Strategy

The impact of conflict risk on infant mortality is investigated using a Difference-in-Differences identification strategy. Treatment and control groups are defined at enumeration area level (household cluster) and all of the variation in childhood exposure to treatment across enumeration areas and birth cohorts is exploited. I first estimate the following equation using ordinary least squares.

$$y_{i(m,t,e)} = \gamma conflict_{i(t,e)} + \mu_e + \beta_t + \theta_{hh} X_m^{hh} + \theta_1 X_{i(m,t,e)}^{(1)} + \theta_2 X_{i(t,e)}^{(2)} + \epsilon_{i(m,t,e)}, \quad (1.1)$$

where i denotes an individual child of cohort (year-month) t, born to mother m who lives in enumeration area e. $y_{i(m,t,e)}$ is a dummy equal 1 if child i dies during the first 12 months of her life. $conflict_{i(t,e)}$ is the measure of violence related stress that child i has been exposed to, in utero or during her first year of life. μ_e and β_t are enumeration area and birth cohort fixed effects. $X_{i(m,t,e)}^{(1)}$ and $X_{i(t,e)}^{(2)}$ are observable characteristics at, respectively, individual level (birth order, age gap with direct older and younger siblings, etc.) and enumeration area level (climatic regressors like temperature and precipitations in the corresponding period and before). X_m^{hh} are household level controls (gender and age of household head, wealth index of household, education of the mother, etc.). Standard errors are clustered at enumeration area level.

In the most basic version of the specification shown in equation (1), I control for enumeration area fixed effects to account for permanent unobserved characteristics of the place of residence and cohort fixed effects to account for cohort specific shocks. The full version of the main specifications adds controls for the relevant mother and child characteristics.

The coefficient γ measures the average difference in changes in the probability of death of born babies, between war and non war areas, holding constant all the other relevant characteristics. The implicit assumption behind the identification strategy is that after controlling for cohort fixed effects, enumeration area and household characteristics (or mother fixed effects), and other relevant exogenous covariates, changes in infant mortality would be similar across war and non-war areas in absence of conflict.

The choice of infant mortality as main outcome variable is due to three main reasons. First, infant death is one of Africa's largest health problems. To this day, close to 10% of children born on the continent die before the age of one. Second, infant mortality is one of the standard proxies for economic development (Kudamatsu, 2012), and it captures well variations in factors such as quality of health care, water quality and food supply. Finally, focusing on infant mortality has methodological advantages: it is available for all the children in retrospective fertility surveys, and the relevant period of potential exposure to conflict is relatively short (in utero or during first year of life) decreasing therefore the likelihood that the estimated impact is driven by other confounding factors.

Interpretation of Coefficient

The coefficient γ does not represent the national impact of conflict on infant mortality but the average effect with respect to local averages, cohort averages and purged of region specific flexible trends. The enumeration area fixed effects are also eliminating some of the valuable cross-sectional variation in conflict exposure.

1.5 Empirical Results

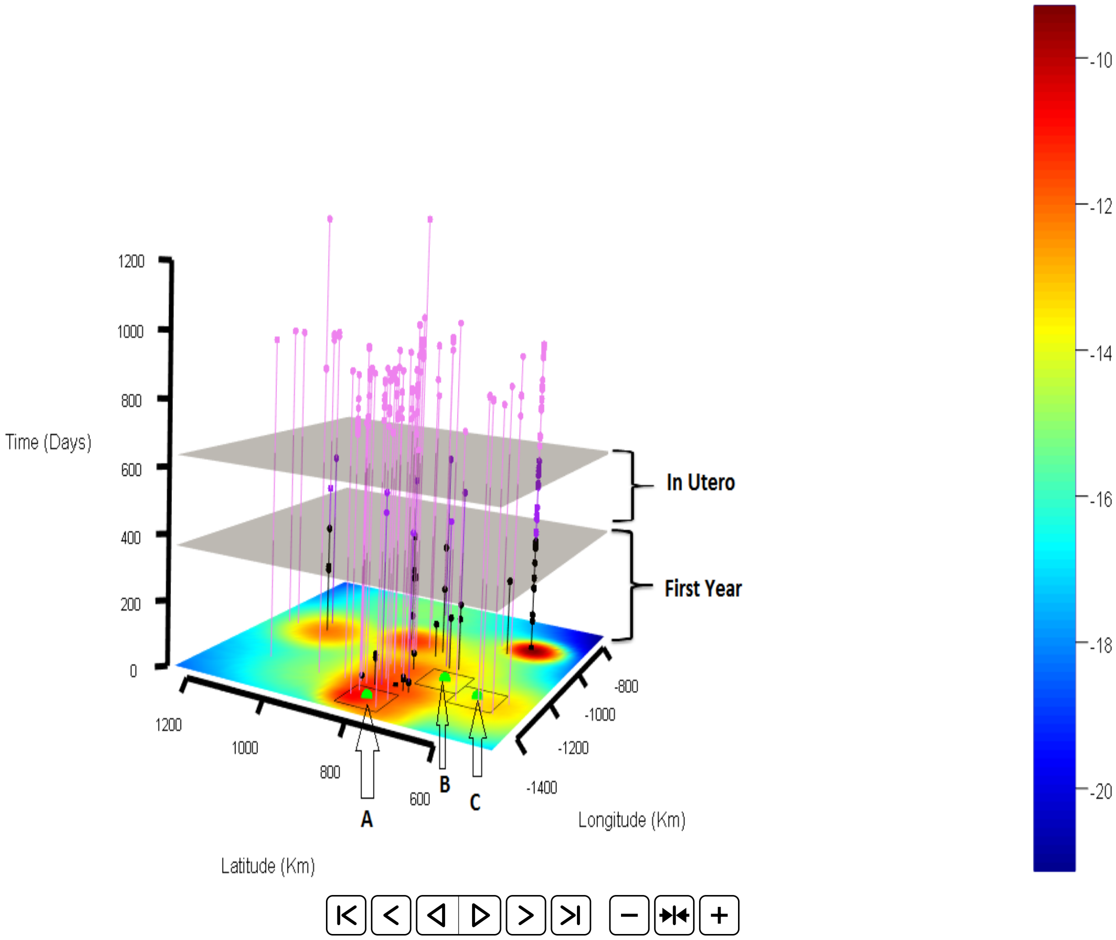
1.5.1 Estimated Conflict Risk

The adaptive kernel density estimation method described above is performed on conflict data for Ivory Coast and Uganda.[16] The risk of exposure to a conflict event between

[16] Average estimated spatial smoothing parameters (h_s) are 34 and 28 Km for Ivory Coast and Uganda respectively. The temporal ones (h_t) are 508 and 207 days respectively.

conception and the first year of life is obtained by integrating the estimated density over the relevant space-time window. As an illustration, figure 1.6 shows the distribution of events in space and time and the probability of being exposed to an event in utero or during the first year of life for children conceived in September 2003 in Ivory Coast. Gray planes delimit the period in utero. Events in purple vertical bars occured in utero, events in black vertical bars occured during first year of life and the violet ones before conception. The heatmap at the bottom shows the likelihood that an event happens in utero or during the first year of life in log scale. For household clusters A, B and C for instance, there was no event happening in the relevant time window (conception to year 1) within 100 Km but the risk of an event was high. Cluster A has experienced many events just before conception and also during the relevant time window but beyond the 100 Km radius. High risk in cluster B is however only driven by events that occured beyond the 100 Km radius while risk in Cluster B is driven by events that happened before the relevant time window. These are examples of situations in which the new approach captures high risk but not the standard approach of counting events or casualties in a given space-time window. I show below that indeed children born in clusters like A, B and C also suffer major health setbacks.

Figure 1.6: Risk for children conceived in September 2003 - Log Scale



Preliminary Observations

Before investigating the impact of violence risk on infant mortality, it's important to discuss the (dis)similarities between the proposed risk measure and the standard metrics used in the literature. Table 1.1 shows the distribution of my sample according to the risk of exposure to violence and the number of observed conflict events (in utero or during the first year of life) within 50 km around the place of birth of each child in Ivory Coast. The probability of being exposed to an event for children born after 1994 is grouped in four quartile groups.

There is some correlation between the proposed measure of violence risk and the standard metrics used in the literature, but the two metrics also differ substantiallly. There was no event happening within 50 km around the place of birth of almost 90% of observations in the two lowest quartiles of the violence risk. This percentage decreases to 60% and 15% respectively for the third and top quartiles of the risk. Observations that experience at least 14 events within 50 km in utero or during their first year of life belong exclusively to the top quartile of the risk distribution. On the other hand, within the group of observations with no event within 50 km, only 36% belong to the lowest quartile of the distribution of violence risk. 35% and 23% of them are in the 2nd and 3rd quartiles of the risk measure respectively and the rest belongs to the top quartile. This leaves room for refining the treatment status for observations that the standard measures of conflict exposure cannot distinguish.

Table 1.1: Distribution of children by risk of exposure to a conflict event and levels of observed violence in Ivory Coast

	Conflict events within 50km					
Risk quartiles	0	[1 ; 2]	[3 ; 4]	[5 ; 13]	14 +	Total
Q1	1,144	114	24	1	0	1,283
Q2	1,118	109	22	6	0	1,255
Q3	723	224	141	141	0	1,229
Q4	183	176	170	274	393	1,196
Total	3,168	623	357	422	393	4,963

Rows represent the probability of exposure to at least one event in utero or during the first year of life. It is grouped in quartiles. Columns represent the number of observed conflict incidents in utero or during the first year of life within 50 km of children's place of birth.

Source of Variation in the Risk Measure

Table 1.2 shows the source of the variation that generates the children exposed to high risk of violence but not experiencing any event within 50 km during the relevant period of time (yellow cells in Table 1.1). It show that 22 % of these observations experienced an event within 50 km the year before their conception or the year after their first year of anniversary (time lags and forwards). By extending the size of the area considered to 150 km and keeping the 1 year time lag before and after the relevant period of time, we cover 90% of these observation. However, occurrence of events in space and time are correlated in such a way that most of these children experience at least one event in spatial lag and one event in temporal lag (75 % overlap).

Table 1.2: Source of Variation in the Risk Measure

	Share of observations with events:			
Area radius (Km)	only in space	only in time	in space and time	Total
50	-	21.51	-	21.51
100	9.53	23.92	37.36	70.81
150	5.62	12.28	74.22	92.12

Table 1.3 shows the distribution of my sample in Uganda according to the risk of exposure to violence and the number of observed conflict events (in utero or during the first year of life) within 50 km around the place of birth of each child. The probability of being exposed to an event is grouped in four quartile groups.

Table 1.3 shows some correlation between the proposed violence risk measure and the standard metrics used in the literature, but it also shows some substantial variations. Children exposed to some events also experienced high risk. However, a big share of children that did not experience any event had a high risk of exposure to violence in utero or during their first year of life.

1.5.2 Violence Risk and Infant Mortality: Main Results and Robustness

Table 1.4 shows the estimated effects of being exposed to high risk of violence on the likelihood of dying within the first 12 months of life in Ivory Coast. Column (1) shows

Table 1.3: Distribution of children by risk of exposure to a conflict event and levels of observed violence in Uganda

	Conflict events within 50 Km					
Risk Quartiles	0	1	[2 ; 3]	[4 ; 12]	13 +	Total
Q1	1,192	0	0	0	0	1,192
Q2	1,227	20	8	0	0	1,255
Q3	1,075	69	51	6	0	1,201
Q4	685	126	156	100	154	1,221
Total	4,179	215	215	106	154	4,869

Rows represent the probability of exposure to at least one event in utero or during the first year of life. It is grouped in quartiles. Columns represent the number of observed conflict incidents in utero or during the first year of life within 50 km of children's place of birth.

the estimated coefficients for the basic specification from equation (1) in which I control only for enumeration area fixed effects and cohort fixed effects. The variable of interest is the standardized estimated level of risk of exposure to an event in utero or during the first year of life. The estimated γ coefficient is positive, significant. This coefficient increases and remains significant after progressively adding controls for child characteristics in column (2) and family characteristics in column (3). The main specification in column (3) suggest that a standard deviation increase in violence risk increases the likelihood of dying by 1 percentage point. Column (4) controls for number of events that happen within 25, 50, 100 and 150 Km and the estimated γ remains stable and significant.

Column (5) splits the observations into 4 groups to test some of the implications of the risk model. The control group are the children that are both exposed to low risk of violence and have not experienced any event within 50 km (blue cell in Table 1.1). Compared to this group, children that are exposed to low risk but experienced some events within 50 km (gray cells in Table 1.1) do not suffer any health setback. Those exposed to high risk of violence without any event (yellow cells in Table 1.1) suffer major health setbacks comparable to those exposed to high risk of violence with some events within 50 km (red cells in Table 1.1). The relevant treatment metrics is therefore in line with the risk model and avoids false positives that come from isolated events that do not generate any fear and true negatives that come from cohorts not exposed to direct violence but born under high risk.

The magnitude of the estimated effect is substantial. The gap in infant mortality rate between children in top and bottom quartets of risk distribution is of 6 percentage points, which represent more than 50% of average infant mortality rate.

Table 1.4: Impact of violence risk on infant mortality in Ivory Coast

	Infant Mortality				
VARIABLES	(1)	(2)	(3)	(4)	(5)
Standardized violence risk	0.008*	0.010**	0.010**	0.010*	
	(0.005)	(0.005)	(0.005)	(0.005)	
Low violence risk with at least 1 event within 50 km					-0.005
					(0.025)
High violence risk with no event within 50 km					0.053**
					(0.021)
High violence risk with at least 1 event within 50 km					0.069***
					(0.023)
Observations	4,963	4,963	4,944	4,944	4,944
R-squared	0.138	0.182	0.184	0.186	0.186
Cohort FE	YES	YES	YES	YES	YES
Enumeration Area FE	YES	YES	YES	YES	YES
Family characteristics	NO	NO	YES	YES	YES
Child characteristics	NO	YES	YES	YES	YES

High violence risk" is a dummy equal 1 if the risk is high enough not to belong to the bottom quartile of the distribution. Full set of controls includes mother's height, education, gender and age of household head, household wealth index, 12 months rainfalls (for the each of the two years preceding the birth of the child and during her first year of life), birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.5 shows some robustness of the baseline specification. Column (1) includes region specific time trends, column (2) clusters the standard errors at 50x 50 Km PRIO-GRID cell level, column (3) uses older cohorts (born before 1995). Column (4) restricts the sample to observation that did not experience any event within 50 Km. The magnitude of the estimated coefficient is stable but standard errors are incresed. Column (5) uses risk measure in log scale. The estimated coefficient suggets that 1 % increase in risk measure increases infant mortality by 0.4 %. The last column shows that the results do not hold for male sample.

Table 1.5: Robustness Ivory Coast

	Girls					Boys
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Standardized violence risk	0.010**	0.010*	0.008*	0.010		0.001
	(0.005)	(0.005)	(0.004)	(0.092)		(0.006)
log of violence risk					0.004**	
					(0.002)	
Observations	4,944	4,944	5,239	3,155	4,944	5,086
R-squared	0.197	0.184	0.179	0.235	0.185	0.195
Cohort FE	YES	YES	YES	YES	YES	YES
Enumeration Area FE	YES	YES	YES	YES	YES	YES
Region specific time trend	YES	NO	NO	NO	NO	NO
Family characteristics	YES	YES	YES	YES	YES	YES
Child characteristics	YES	YES	YES	YES	YES	YES

Full set of controls includes mother's height, education, gender and age of household head, household wealth index, 12 months rainfalls (for the each of the two years preceding the birth of the child and during her first year of life), birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.6 is the equivalent of Table 1.4 for Uganda. The qualitative results are pretty similar. A standard deviation increase in violence risk increases infant mortality by 0.8 percentage points in Uganda. Children exposed to high risk of violence suffer major health setbacks even when the risk does not realize into actual event. Table 1.7 is the equivalent of Table 1.5 for Uganda.

Table 1.6: Impact of violence risk on infant mortality in Uganda

	Infant Mortality				
VARIABLES	(1)	(2)	(3)	(4)	(5)
Standardized violence risk	0.011***	0.009**	0.008**	0.008**	
	(0.004)	(0.004)	(0.004)	(0.004)	
High risk with no event within 50 km					0.018**
					(0.008)
High risk with at least one event within 50 km					0.024*
					(0.014)
Observations	4,869	4,869	4,868	4,868	4,868
R-squared	0.109	0.155	0.160	0.161	0.160
Cohort FE	YES	YES	YES	YES	YES
Location FE	YES	YES	YES	YES	YES
Child characteristics	NO	YES	YES	YES	YES
Family characteristics	NO	NO	YES	YES	YES

High violence risk" is a dummy equal 1 if the risk is high enough not to belong to the bottom quartile of the distribution. Full set of controls includes mother's height, education, gender and age of household head, household wealth index, 12 months rainfalls (for the each of the two years preceding the birth of the child and during her first year of life), birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.7: Robustness Uganda

	Girls					Boys
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Standardized violence risk	0.008**	0.008**	0.009*	0.011		0.001
	(0.004)	(0.004)	(0.005)	(0.049)		(0.003)
log of violence risk					0.002	
					(0.002)	
Observations	4,868	4,868	2,463	4,178	4,868	5,060
R-squared	0.161	0.160	0.202	0.171	0.160	0.188
Cohort FE	YES	YES	YES	YES	YES	YES
Enumeration Area FE	YES	YES	YES	YES	YES	YES
Region specific time trend	YES	NO	NO	NO	NO	NO
Family characteristics	YES	YES	YES	YES	YES	YES
Child characteristics	YES	YES	YES	YES	YES	YES

Full set of controls includes mother's height, education, gender and age of household head, household wealth index, 12 months rainfalls (for the each of the two years preceding the birth of the child and during her first year of life), birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

1.5.3 Placebo Test

Table 1.8 shows some placebo that test whether the results are driven by pre-existing trends. In column (1) and (3), I randomly assign risk values within each enumeration area for Ivory Coast and Uganda respectively. The estimated coefficients are close to zero and statistically insignificant. In column (2), I run a placebo test for Ivory Coast in which children born in 1995 and 1996 are used as fake war cohorts and treatment intensity is given by the maximum risk in a given area over time.[17] column (4) does a similar exercise for Uganda using children born between 2006 and 2011 as fake war cohorts.[18] Results confirm that there is no pre-existing trend difference between areas with high and low violence risk.

Table 1.8: Placebo test and alternative spefications

	Ivory Coast		Uganda	
VARIABLES	(1)	(2)	(3)	(4)
Standardized violence risk with random values	0.001		0.004	
	(0.004)		(0.004)	
Maximum standardized violence risk		0.000		
X 1995-1996 cohorts		(0.011)		
Maximum standardized violence risk				-0.004
X 2011-2016 cohorts				(0.003)
Observations	4,944	1,621	4,868	3,739
R-squared	0.189	0.324	0.160	0.163
Cohort FE	YES	YES	YES	YES
Enumeration Area FE	YES	YES	YES	YES
Family characteristics	YES	YES	YES	YES
Child characteristics	YES	YES	YES	YES

Full set of controls includes mother's height, education, gender and age of household head, household wealth index, 12 months rainfalls (for the each of the two years preceding the birth of the child and during her first year of life), birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

[17] Children born before 1995 are the non exposed cohorts in column (2) of Table 1.8.

[18] Children born after 2011 are the non exposed cohorts in column (4) of Table 1.8.

1.5.4 Potential biases

The causal interpretation of the estimated coefficient can be threatened by factors such as omitted variable bias, selective migration and fertility. This section presents evidence that the main results found here are unlikely to be driven by such factors. The magnitudes of the estimated coefficients are, if anything, lower bounds of the true impact of conflict risk on infant mortality.

Omitted Variable Bias

A possible threat to the Difference-in-Differences identification is the existence of time varying omitted factors that drive the timing and location of violence and, at the same time, are correlated with infant health. I show in this section that this is not an issue for the empirical results presented above.

For the case of Ivory Coast, figure 1.1a shows that violence intensity varies drastically over time according to the political agenda, both increasing and decreasing sharply, so that contemporary changes in omitted determinants of health are less likely to be driving the estimated coefficients.[19]

One of the main drivers of conflict intensity in Sub Saharan Africa is the value of lootable resources like valuable minerals (Berman et al., 2017; Dagnelie et al., 2018). An increase in prices of minerals can also affect child health through household income for instance, so it is worth exploring this channel specifically in theory. However, the context of the Ivorian conflict rules out such option. Lootable resources play a minor role in the Ivorian economy and there is no anecdotal evidence linking them to the conflict under scrutiny. The country's economy relies mostly on cash crops that require permanent and heavy infrastructure to be sent to the world market like cocoa (first world producer), coffee, raw cashew nuts, palm oil, etc.

The identification strategy in Uganda relies on the exogenous increase in violence intensity in Northern Uganda following the change in the policy against internal insurgency that occurred in 2001, after the September 11 terrorist attack. Violence in Uganda be-

[19] The variations in violence intensity are even more drastic at local level. The space and time variation in violence intensity also helps in the identification strategy. Treatment (being exposed to high risk of violence) can happen to children of any birth order within the same family and at different periods in time, mitigating even more the effect of potential confounding factors.

tween 2002 and 2005 happened as a consequence of this exogenous shock and led to the ending of any significant activity from the rebel groups that were active in the country.

Migration

Conflict induced displacements could bias downward the estimated impact of conflict on infant mortality since I do not observe migration history in my data. However, permanent migration is rare in countries like Ivory Coast and Uganda. In rural areas for instance, land is the most valuable asset and the absence of property right makes it risky for farmers to leave their lands unused for too long or to try to establish themselves in another locality. IDMC (Internal Displacement Monitoring Center) estimated that, as of February 2015, over 80 % of the 2.3 million people displaced by violence since 2002 in Ivory Coast had managed to return to their homes. Moreover, conflict induced temporary displacement is one of the channels through which conflict affects infant mortality. Displaced households often live in camps or host communities with limited access to basic needs like health services, clean water or the ability to undertake an economic activity.

Fertility

Selective fertility decisions could be a threat to the identification strategy. If the fear of being exposed to conflict affects the fertility decisions of mothers differently depending on some specific mother/household characteristics, then any estimate of the consequences of the shock would be biased. One would expect households with high socio-economic status to be most likely to adjust their fertility decisions to violence risk.[20] The stability of the estimated effect when controlling for household and mother characteristics is a first signal that endogenous fertility is not an issue here.

In the case of Ivory Coast for instance, the empirical evidence suggests that households who had at least one child exposed to violence in utero or during their first year of life are similar to the other households living in conflict affected areas as shown in Appendix Table A2.[21]

[20] High socioeconomic status households are more likely to have different preferences due to their education. They tend to have less children and invest more in them.

[21] I examined characteristics like mother's education, number of births, height, as well as household head's age and gender. The only significant difference in balance Appendix Table A2 is the number of children born in treated versus non-treated households in rural areas. The treated families have more children but this could be explained by the fact that the likelihood of having at least one child treated increases with the number of children that you have.

Many factors could explain the fact that it is unlikely to observe selective fertility behaviors with respect to violence in countries affected by low intensity conflicts. In general, developing countries in Sub-Saharan Africa are facing high infant and child mortality rates. This high probability of loosing a child combined with the low cost of having an extra one and deeply rooted pro-natalist religious and cultural beliefs have led to high birth rates and a somehow fatalistic view over the survival of children.

1.5.5 Comparison with Existing Measures of Exposure to Violence

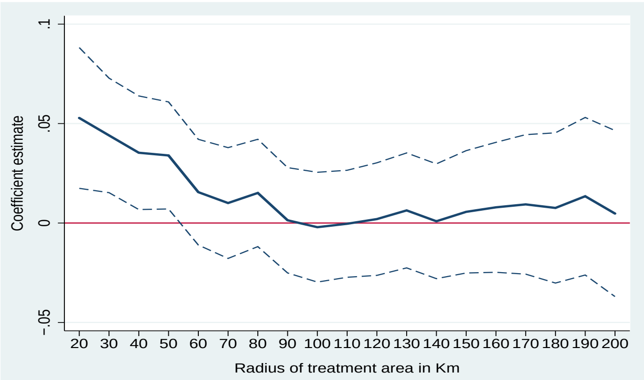
In theory, one of the benefit of the proposed measure of violence risk over the standard metrics used in the literature is its ability to discriminate between children that are exposed to high and low risk of violence even at equal level of (non) exposure to observed events. This section shows that this benefit is also empirically relevant.

First, I show the estimated impact of conflict on infant mortality using the standard metrics in the literature. Figure 1.7 plots several values of the coefficient γ estimated from the main specification in equation (1) for Ivory Coast. The conflict variable is defined as a dummy equal 1 if there was at least one event that happened within a certain radius r from the place of birth of a given child. A new regression is ran for each tick in the x axis and the coefficient together with 95% confidence bands for γ are plotted and interpolated to give the continuous lines. The first regression considers a treatment area of 20 km of radius around the household location (large enough to encompass a city).[22] Around 20% of the observations are classified as treated and their likelihood of dying before turning one year old increases by 5 percentage points. The point estimate is significant and quite substantial. It represents a 50% increase in the average infant mortality rate. As the radius increases, we are moving more observations from the control to the treatment group and the magnitude of the coefficient declines towards zero. It is however still substantial and significant up to 50 km (big enough to encompass

[22] In DHS data, some noise is introduced in the GPS coordinates of the enumeration areas that are surveyed for privacy reasons. They are displaced by up to 2 km in urban areas and 10 km in rural ones. For this reason, I use at least 20 km to define geographic areas that contain for sure the surveyed households.

a department - 2nd/3rd administrative level). Beyond this size, the coefficient is small and not significant meaning that we start putting in the treatment group observations that are not treated and these observations bring the estimated difference down to zero.

Figure 1.7: Impact of conflict on infant mortality using observed violence



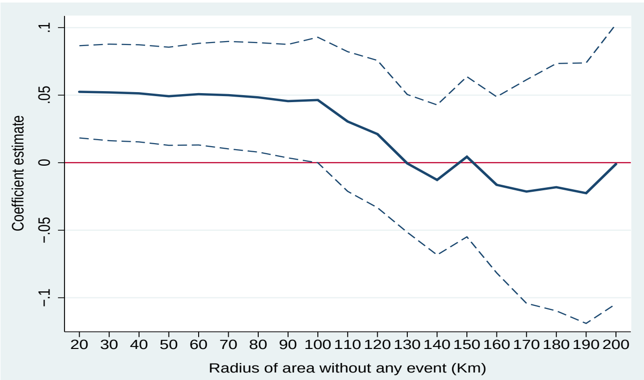
Notes: Estimated impact of conflict on infant mortality using the main specification in equation (1). The $conflict$ variable is defined as a dummy equal 1 if there was at least one event that has happened within a certain radius r from the place of birth of each child (in utero or during first year of life). A new regression is ran for each tick in the x axis and the γ coefficients together with 95% confidence bands are plotted and interpolated to give the continuous lines. Robust standard errors are clustered at enumeration area level. The first regression considers a treatment area of 20 km of radius around the household location, the second 30 km, etc.

This sensitivity of the magnitude of the estimated impact to the size of the area considered is well documented in the literature. Researchers usually settle for the specification with the smallest standard errors and show some robustness to increasing or decreasing this size.

Figure 1.8 shows the estimated impact of conflict using the proposed risk measure and restricting the analysis to observations with no event within r km from their place of birth. These observations are not treated according to the standard metrics used in the literature and the question here is whether the proposed risk measure is relevant enough to capture significant variations in the risk of exposure to violence within this sample. Small values of r mean more observations to split between treatment and control. The main specification is the regression that considers observations with no event within 50 km (first column of Table 1.1). The estimated coefficient of the impact of being exposed to high risk of violence is strikingly constant (increase of 5 percentage points in infant mortality) and falls apart only when the sub-sample is too small because we are looking at cohorts with no event within more than 150 km around their area of birth. The differences in the risk measure are just noisy in this sub-sample because they are all barely affected by the violence risk. This also explains the large confidence bands after 150 km in the graph.

Results in Figure 1.8 imply that the risk measure proposed in this paper is able to split an otherwise homogeneous control group (according to the classic methods used in the literature) into treatment and control group in such a relevant way that the magnitude and significance of the estimated effect is similar to comparing cohorts with observed events happening in their neighborhoods to the others.

Figure 1.8: Impact of violence risk on infant mortality: sub-sample of children with no event within r km



Notes: Estimated impact of violence risk on infant mortality when restricting the sample to non treated observations according to standard metrics in literature. The variable of interest is a dummy equal 1 if the risk of exposure to violence is high (belongs to the second quartile or higher). The sample used for each regression is restricted to observations that had no event happening within a certain radius r from the place of birth (in utero or during first year of life). A new regression is ran for each tick in the x axis and the γ coefficients together with 95% confidence bands are plotted and interpolated to give the continuous lines. Robust standard errors are clustered at enumeration area level. The first regression considers a treatment area of 20 km of radius around the household location, the second 30 km, etc.

1.5.6 Analysis of Transmission Channels

Understanding the specific mechanisms by which violence risk impacts child health is critical for developing adequate policy responses to protect children from this adverse effect. The timing of violence escalation and DHS survey years in Uganda allows me to partially investigate these channels. The DHS surveys provide indeed a rich set of information on the use of health services, child and maternal health for pregnancies that happened at most 5 years before the survey year. Using this information, I focus on exposure to violence risk in utero and show how it affected key health inputs and outcomes.

Column (1) of Table 1.9 shows first that exposure to high risk of violence in utero increases significantly infant mortality. Column (2) and (3) suggest that violence risk decreases the likelihood of being born with low birth weight. High violence risk also decreases the likelihood of receiving prenatal care during the first quarter of pregnancy (column (4)), delivering in a health center (column (5)) or having access to a C-section (column (6)). Women exposed to high risk of violence during a specific pregnancy also have longer duration of postpartum amenorrhea which is evidence of maternal stress and insufficient nutrition.

These results suggest that conflict in Uganda affected the quality of health services and their use by citizens during crucial times. It also induced maternal stress and short term malnutrition.

Table 1.9: Channels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Infant Mortality	BW<=2.5 Kg	Small at birth	Q1 Prenatal care	Home delivery	C section	Duration Amenorrhea
Standardized violence risk in utero	0.008**	0.016*	0.008	-0.010**	0.014**	-0.015***	
	(0.004)	(0.009)	(0.009)	(0.005)	(0.007)	(0.005)	
Standardized violence risk in utero and during 1st year							0.181**
							(0.077)
Mean depedent variable	0.044	0.096	0.239	0.158	0.376	0.087	8.166
Quantification: Gap Q4-Q1	0.030*	0.020	0.036*	-0.032	0.054*	-0.025	0.737*
	(0.015)	(0.014)	(0.020)	(0.021)	(0.028)	(0.026)	(0.430)
Observations	4,868	2,673	4,785	4,868	4,868	3,025	4,838
R-squared	0.160	0.123	0.085	0.195	0.259	0.128	0.264
Cohort FE	YES	YES	YES	YES	YES	YES	YES
Enumeration Area FE	YES	YES	YES	YES	YES	YES	YES
Family characteristics	YES	YES	YES	YES	YES	YES	YES
Child characteristics	YES	YES	YES	YES	YES	YES	YES

Full set of controls includes mother's height, education, gender and age of household head, household wealth index, 12 months rainfalls (for the each of the two years preceding the birth of the child), birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

1.6 Conclusion

This chapter analyzed the impact of violence risk on child health using data from Ivory Coast and Uganda. It is based on the intuition that economic agents often react to the risk of being exposed to violence in a given space-time window even before/without any manifestation of violence in this window. These behavioral responses to insecurity can lead to disruptions in the supply and demand of goods and services and affect affect households. A new metric that captures perceived violence risk on the ground is therefore proposed in the paper and used to evaluate the impact of violence risk on child health in a Difference-in-Differences setting. The proposed measure of violence risk is based on an estimation of the underlying distribution of the violence process in space and time.

The empirical results show that cohorts of children exposed to high risk of violence suffer major health setbacks. In particular, it increases infant mortality by more than 50% in both Ivory Coast and Uganda. This effect is similar in magnitude and significance level even when the violence does not manifest itself in the relevant space-time window. These results suggest that conflict is a public bad that affects entire communities through their risk coping behaviors. An investigation into the potential channels through which this effect operates suggests that we cannot rule out factors like maternal stress, malnutrition and deterioration of quality/decrease in use of health services.

This paper has important policy implications for different stages of violent conflicts. First, We should focus even more on conflict prevention efforts to avoid the huge humanitarian and economic cost that will arise if violence breaks out. Second, during conflict, governments and NGOs should address fears/expectations of economic agents and prevent disruptions in supply and demand of goods and services in order to minimize the cost of ongoing threats. Finally, in post conflict reconstruction settings, the findings in this paper imply that policy interventions should include all the children born under violence stress and not just the direct victims of violence.

Chapter 2

Locust Plagues and Child Health in the Sahel Region

written jointly with Bruno Conte Leite and Lavinia Piemontese.

2.1 Introduction

Conditions experienced early in life have long-lasting effects on various socioeconomic outcomes. In particular, it is well established that harsh conditions experienced in utero can have detrimental and persistent effects on individual health throughout the whole life cycle (Stein et al., 1975). This concept, known as the foetal origins hypothesis, implies that it is harder to remedy bad fetal health later on in life.[1] The empirical literature has successfully documented the impact of adverse in utero conditions on infant mortality (Kudamatsu et al., 2012; Dagnelie et al., 2018), child health (Akresh et al., 2011, 2012a; Bundervoet et al., 2009), adult health (Akresh et al., 2012b; Maccini and Yang, 2009), and adult socioeconomic outcomes such as literacy, educational attainment, income and labour market status among others (Almond et al., 2007; Alderman et al., 2006; Lavy et al., 2016).

This growing literature has focused mostly on poor countries in Africa, where child

[1] Almond and Currie (2011) provide an excellent review on the foetal hypothesis theory and its theoretical relation to later life outcomes.

malnutrition and human development are still lagging behind international standards. They have relied heavily on the exposure to exogenous shocks in utero to identify their causal impact on later outcomes. Most of these shocks have been in the form of weather/ agricultural shocks, exposure to civil conflicts and/or changes in institutional setup (i.e. ethnic favouritism, Kudamatsu (2009)). One particular type of agricultural shock that recurrently puts at risk food supply in many developing countries are desert locust plagues. Desert locusts are grasshoppers that live around the Sahara desert. Under favorable breeding conditions they multiply, gregarise, change in morphology (color) and become more voracious. At the stage of a plague, they fly in large swarms, take breaks to rest, feed and start moving again leaving behind destroyed vegetal cover. Little attention has been devoted to understanding the consequences of this agricultural shock.[2]

This paper aims at filling this gap in the empirical literature by studying how in utero exposure to locust plagues affects child health. We specifically study the impact of the 2004 locust plague invasions in Mali and Senegal. We took advantage of the local nature of the plague to identify causal effects of the exposure to the plague on child health. Our identification strategy relies on spatial variation in exposure of different birth cohorts in both countries.

Using a Difference-in-Differences identification strategy, we show that children exposed in utero to the adverse effects of locust plagues suffer major health setback in Mali and Senegal. Exposed children have, on average, a Z-score 0.24 points and 0.48 points lower than non-exposed children in Mali and Senegal respectively. The estimated effect is stronger for poorer households. Children exposed in utero to the speculative price effect suffer as much as those exposed to the crop failure effect in utero. We find no effect for cohorts of children exposed to the shock after birth.

This paper contributes to two main strands of the economics literature. First, it contributes to the previously discussed literature on the importance of early life conditions (Lavy et al., 2016; Maluccio et al., 2009; Black et al., 2007; Behrman and Rosenzweig, 2004; Stein et al., 1975). A substancial part of this literature has been studying the effect on child health of exposure to weather shocks (Maccini and Yang, 2009), civil wars

[2]An exception is a work by De Vreyer et al. (2014) that estimates the effect of exposure to the 1987 locust plague in Mali on educational outcomes.

(Dagnelie et al., 2018; Valente, 2015) or adverse institutional setup (Kudamatsu, 2012). We complement this literature by investigating the impact of locust swarm invasions on child health. We show that this shock has long lasting effects on children exposed in utero but not after birth. We also show that the adverse effects of this shock are present for both male and female children and the effect is stronger in poorer households.

Second, this paper is also related to the literature that studies the consequences of negative agricultural shocks caused by pest invasions. Baker et al. (2018) show that the boll weevil's pest invasion that affected US cotton production from 1892 to 1922 led to an increase in educational attainment due to reduced opportunity cost for schooling. Banerjee et al. (2010) show that phylloxera invasion in 19th century France affected wine production and led to substantial effects on adult height for people born in affected areas in that period. De Vreyer et al. (2014) used locust plague invasion in 1987-1989 in Mali to show that it had a long term effect on educational attainment. Our paper is closer to the latter one since we study the consequences of the same pest species. We depart from this paper by looking at how it affects child health and we use a more recent plague episode that has affected the Sahel region. Our main contribution to this literature comes from the fact that we are the first ones, to the best of our knowledge, to make the distinction between the speculative price effect and the crop supply effect of pest invasions that affect agricultural production.

The remainder of this chapter is organized as follows. Section 2.2 provides some background on the 2003-05 locust plague in Western Africa and its relation to food shortages and child health in Mali and Senegal. Section 2.3 presents the data used for the empirical estimations and Section 2.4 discusses the channels through which locust invasions can affect the well-being of households. Section 2.5 presents the empirical strategy used and Section 2.6 shows the results obtained. Section 2.7 concludes.

2.2 Background: Desert Locust Plagues

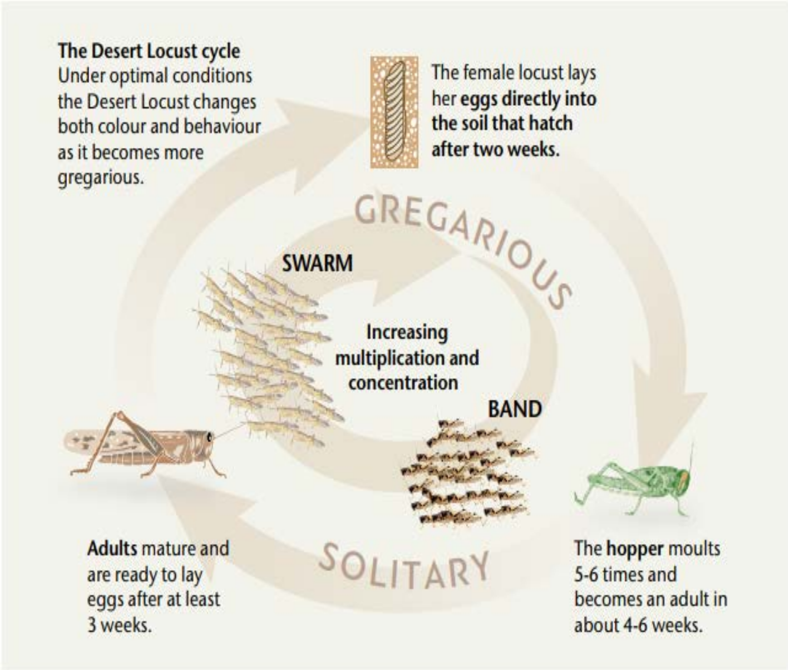
Desert locusts are insects that inhabit the arid and hyper-arid zones of the Sahel region, northern Africa, Middle East and southeast Asia. They pertain to the family of *grasshoppers* and normally inhabit more desertic zones – called *recession areas* – in a solitary, harmless and integrated way with the local ecology. What makes them differ from traditional grasshoppers is their capacity to mutate physically and change behavior under favorable breeding conditions. This happens in periods of excessive rainfall followed by relatively mild temperatures in the *breeding areas* (a subset of the recession areas) leading to faster reproduction cycles. The high density of locust combined with a relative shortage of vegetation leads to a *gregarious phase*: the insects start to mutate physically and behave as a whole group, known as a *swarm.*[3] Upon gregarisation, these groups become more voracious and reproduce faster. If needed, locust swarms will fly away to find a location with vegetation for feeding and suitable conditions for reproducing. Within a few weeks after settling in a new location satisfying those conditions, there can be a new generations of gregarious locusts. The increase in the size of the swarm can be remarkable – the incorporation of locusts from the new generation can lead up to a tenfold increase in the size of the swarm (FAO, 2004). Under favorable breeding conditions and in the absence of (human) controls for a few months, the swarms can become extremely numerous and exceed billions of locusts.

What makes the desert locusts so concerning is that, during the migration process, the swarms roughly follow strong winds that move them far away from the most desertic areas of the Sahara into the central Sahelian and tropical areas in the South or the Mediterranean areas in the North. These are known as *invasion areas*, which span over more than 50 countries with about 29 million square km (Herok et al., 1995). These invasion areas are usually more densely populated and used for agricultural production. Crop yields in such locations can therefore be partially or totally consumed by the pest threatening food supply for entire regions. The same can happen to pasture vegetation used to feed livestock. Moreover, locust swarms can fly for very long distances, over hundreds of kilometers in a single day, so good conditions for gregarisation can easily have consequences for relatively far away locations. Figure 2.2 illustrates the geographical location of the breeding, recession and invasion areas for desert locust.

A subsequent concern is the possibility that climate conditions remain "ideal" for a long time within big geographic areas so that the swarming, breeding and migration

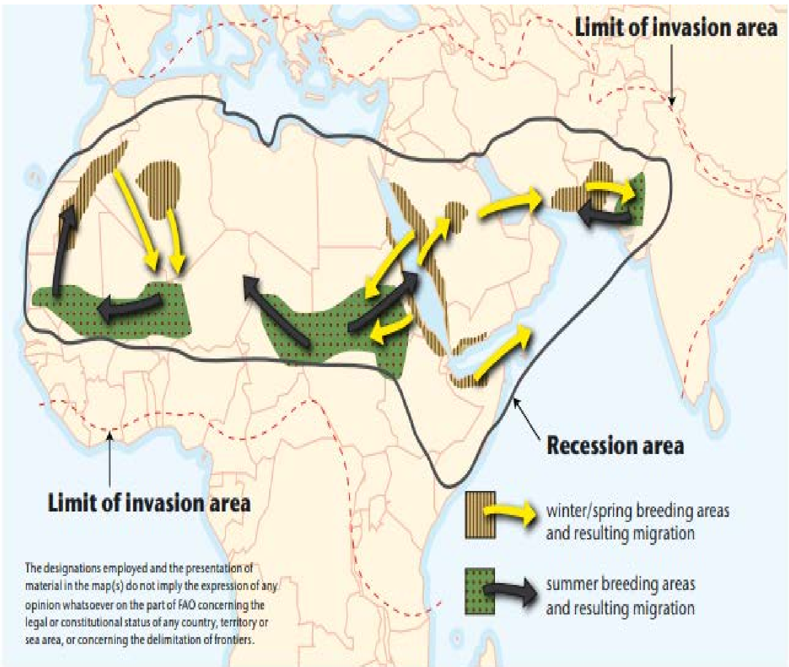
[3] Technically, those groups are named as band if composed of wingless nymphs and as swarm for adult locusts

Figure 2.1: Desert Locust Life Cycle



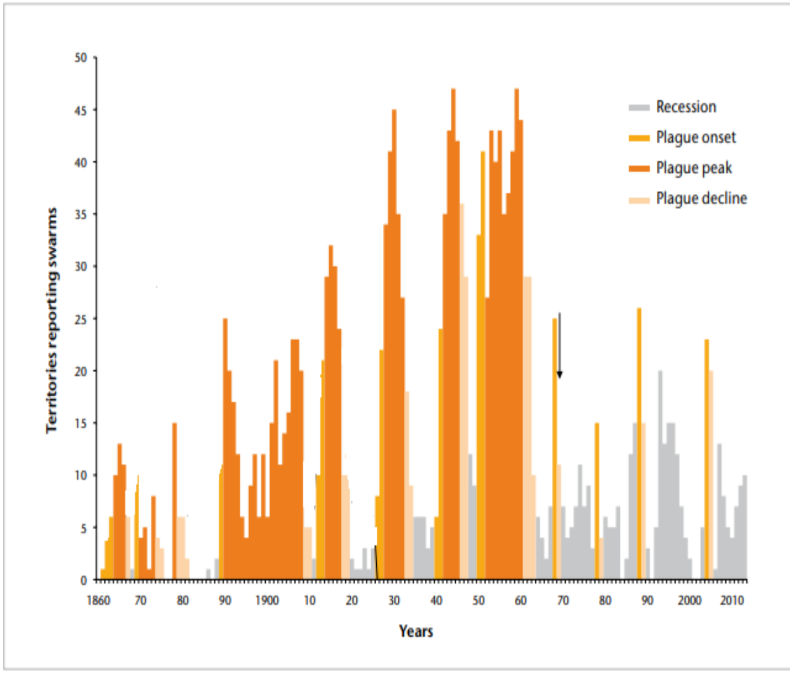
Source: Cressman and Stefanski (2016)

Figure 2.2: Affected Areas



Source: Cressman and Stefanski (2016)

Figure 2.3: Historical plagues



Source: Cressman and Stefanski (2016)

process turns into a regional plague. That is when locust swarms multiply exponentially in size and number and spread over several countries. The consequences can be nefast: a locust swarm can have from one to several hundred square km; a square km size swarm has at least 40 million locusts which can consume daily the equivalent food consumption of 35,000 people (Symmons et al., 2001).

Case of Mali and Senegal

Mali and Senegal are located in a very risky area in terms of likelihood of experiencing locust plague invasions. They are the only two countries in the Sahel region affected by the plague of 2003-2005 with matching household survey data for our analysis. Mali is a landlocked country in West Africa with a surface of about 1.2 million square km, spread in latitude across three different climate zones: the desertic zone (Saharan), the transition zone (Sahel) and the tropical zone. A substantial part of its territory is a locust breeding area, which makes it very likely to experience (or be invaded by) locust swarm outbreaks (see Figure 2.2). Furthermore, according to World Bank Data, Mali is a low-income country (ranked 213^{th} in per-capita terms, with about 50% of the population below the poverty line) whose population, of about 18 million individuals, is largely present in rural areas (about 60%) and sparsely distributed within the country (about 14 people per sq. km, ranked 241^{th}). The country is strongly dependent on agriculture, which accounts for about 60% of its GDP (De Vreyer et al., 2014) and is still largely composed by small-scale subsistence farming. As such, a large share of its population is vulnerable to agricultural hazards such as droughts and plagues.

Not surprisingly, Mali was one of the most damaged countries in the last two plagues in 1987-89 and 2003-05 in that area. The latter one is the focus of this study. It started from optimal climate conditions in late 2003 and escalated to a massive plague throughout the entire Western Africa (Ceccato et al., 2007). The most affected countries were Mauritania, Morrocco, Mali, Niger and Senegal (see Figure B1 and Figure B3). Precise information on the overall impact of this plague is not available from reliable data, but anecdotal evidence points towards harvest losses valued over billions of US$, despite international plague control efforts that cost hundreds of millions of US$. By 2005, multiple countries in the region were facing concerning scenarios of a food crisis, with a large share of agricultural lands and crops damaged. Overall, a grain surplus was registered in the region on average, but this did not prevent affected areas to experience increases in food prices (FAO, 2005a) and food shortages. Reliable estimates of the agricultural losses are not available, but the existing evidence suggests that the plague had devastating effects in Mali and Senegal. By mid-2005, FAO (2005b) reported alarming increases in cereal prices and a deterioration of conditions for livestock production in both countries. Moreover, the poorest households suffered the most from this situation as the report states that by the middle of 2005, "*access to main food staples [was] increasingly difficult for vulnerable households and pastoralists. Severe child malnutrition [was] increasing rapidly. Reports from Mali's Kidal region show[ed] that one-third of children under the age of three [were] suffering from severe malnutrition*" (FAO, 2005b).

2.3 Data

In order to understand how child health was affected by exposure to locust plagues in Mali and Senegal, data from several sources was collected. Three particular sources are highlighted. First, the data from locust swarms were obtained from a rich database of worldwide locust monitoring. Second, the data on children anthropometrics was built from widely used surveys in demographics and health. Finally, geographical and socioeconomic characteristics at the spatial cell level were gathered to complement the two main data sources.

Locust Swarm Data: We collected geographical and temporal incidence of locust swarms from the SWARMS (Cressman, 1997) database. It contains historical, geocoded information on many "locust parameters". Those include local breeding conditions, the incidence of locusts adults, hopper bands, swarms, and many others. For this application, we used their data on locust swarms. The SWARMS system is held and maintained by the FAO Desert Locust Information Service (DLIS) Unit. This unit is monitoring, preventing and controlling locust incidence in the Sub Saharan Africa for over 60 years. For that, a national office at each of the countries in the region conducts field activities. These activities include field incursions into areas prone to locust incidence and reproduction to search for (and code if found) locust bands and swarms. The coding is done in-field with a satellite-based technology; the information is automatically sent to the FAO-DLIS headquarters to be further cleaned if necessary. The data is complemented with information from local villages which self-report to field officers and/or the national DLIS office.

The data on the incidence of locust swarms spans from 1985 to present and covers 43 countries. Figure B2 shows the distribution of locust swarm incidence over time. It is quite consistent with the historical evidence of locust plagues, i.e. there are peaks of events in those years. Analogously, the geographical distribution of the swarms is also consistent. As seen in Figure B1, most of the events from the 1987-89 and 2003-05 plagues are concentrated in the Western/North-Western Africa, as expected. The same holds for the events from the 1992-1995 upsurges in the Horn of Africa and South Asia, and the most recent plagues in the Horn of Africa. Subsequently, that data was matched to the growing season of the main harvest crop (see geographical data below) in the $.5 \times .5$ degrees cell where the locusts swarms were spotted. With that, a dummy of whether the swarm was spotted during the local growing season was constructed. The final dataset is a georeferenced, monthly pooled cross-section of locust swarms.

The consistency of the SWARMS data in the temporal and spatial dimensions does not rule out the possibility of measurement error. First, the potential incapacity of the DLIS National Offices to keep field operations consistently and equally through the whole period and the national territory could lead to miscoding and lack of data in specific areas and/or time periods. Furthermore, many countries in the Sub Saharan Africa experienced periods of political instability and/or civil wars during the period covered by the data. That could compromise the operational capacity of field teams. We ruled out these concerns in our study for the following reasons. First, for our countries of interest (Mali and Senegal) at the time period of our analysis (early 2000's), the National Desert Locust Offices have not faced drastic changes and other difficulties that could affect its field operations.[4] In terms of political instability that could hinder in-field data collection, the only historical evidence is the crisis in Mali since 2012.

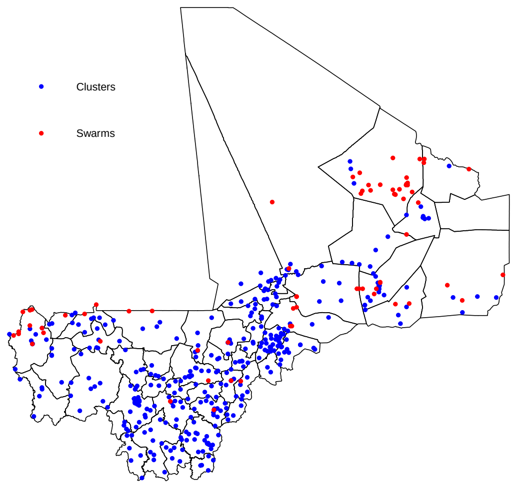
Children Anthropometrics Data: Child height data in Mali and Senegal were obtained from the Demographic and Health (DHS) Surveys. DHS surveys are being conducted in many developing countries since late 1990's with a standardized and nationally representative data collection methodology. The interviews target female individuals, aged from 15 to 49 years old. For some respondents' children aged up to 5 years old, anthropometrics as height and weight are collected by trained surveyors. Importantly, children height are provided in Z-scores, i.e. the difference between the child's height and the mean height of the same-aged international reference population, divided by the standard deviation of the reference population. It provides, moreover, GPS coordinates of the household location, called cluster.[5]

We used for this work the waves V and VI of DHS survey for Mali and Senegal, whose interviews took place in 2006 and 2012/13 for Mali and in 2005 and 2010 for Senegal. As

[4]This information was obtained unofficially with an Information Officer at the DLIS office at FAO, Rome.

[5]Empirical work using DHS data also call the clusters as enumeration areas or locations. We stuck to the terminology of DHS cluster. The clusters' coordinates are released with a random noise for privacy reasons. The noise is set to up to 2 km for household located in urban areas and 5 km for those in rural areas. From the latter, 1% is provided with 10 km noise in the coordinates.

Figure 2.4: 2004 Locust Invasion and Household Clusters in Mali

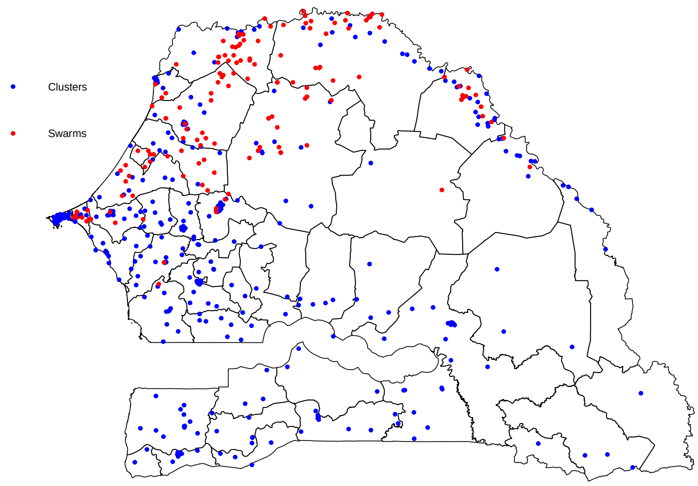


such, the children measured in the in the interviews belong to birth cohorts from 2001 to 2013, covering the period of the 2003-5 plague. In particular, children measured in the first wave are those whose cohort of birth can potentially coincide with the plague period. Children from the second wave, instead, are too young to have been exposed.

To enrich the set of observables available, the data of each child was linked to its household and mother's characteristics. Moreover, it was matched over geographical and socioeconomic characteristics from the spatial cell level dataset using the coordinates of the cluster in which interviews took place. The resulting dataset is a cross-section of children born from 2001 to 2013 with many associated characteristics.

Spatial Cell Data: This paper uses geographical and socioeconomic data from the PRIO-GRID (Tollefsen et al., 2012). It consists of a geocoded dataset, aggregated at spatial cells of 0.5×0.5 degrees. We focused on the statistical information available in that grid, in particular, the local main crop, its growing season, accessibility information and further geographical information. The georeferenced feature of the children and locust swarms dataset allowed us to link information from the grid to the latter, adding further observables to be exploited in our analysis.

Figure 2.5: 2004 Locust Invasion and Household Clusters in Senegal



2.4 Desert Locust Swarms and Household Well-being: Speculative Price Effect and Crop Supply/Income Shock Effect

Figure 2.6 shows the seasonal calendar in Mali and Senegal. It starts with soil preparation between March and May followed by planting in May-June. The growing season happens between July and September and harvest period is in October/November. The 2003-2005 locust plague affected Mali and Senegal during the summer of 2004. The first swarms invaded the North of Senegal on 20th of June 2004 and affected substantially the planting and growing season of the 2004 agricultural season.

Speculative Price Effect

In this paper, we are interested in how these swarm invasions have affected households, specially children in utero which are the most vulnerable household members. Between July and October 2004, households and markets were still relying on harvest of the previous season (2003). There was therefore no local crop supply/income shock in this time window. There could however be a speculative price effect in anticipation of an imminent bad harvest. The mere occurrence of locust swarms could therefore lead to local inflation of prices and harm households with limited resources specially since it happens during the lean season.[6]

Local Crop Supply/Income Shock Effect

The new harvest that is affected by the locust swarms happens between October and November 2004. Households and local markets rely on this potentially bad harvest until the next one in October 2005. We can therefore expect areas affected by locust plague to be treated during all this period by the actual local crop supply shock. This agricultural production shock can affect all vulnerable households through increase in local crop prices. It is moreover an income shock for farmers that had their fields invaded.[7]

Potential Long Term Effect on Agricultural Production

The adverse effects of the plague should disappear therefore only after the harvest of November 2005 unless the crop failure in 2004 affects the quality of seeds used for the next harvest or the income shock led to a depletion of productive assets (livestock for instance). This will lead to a persistent effect (in medium run) on household agricultural production.[8]

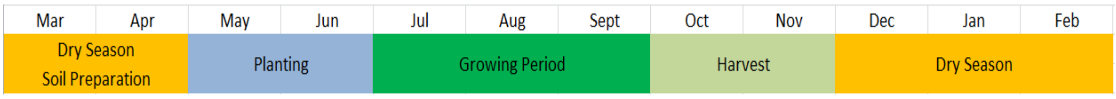
The data at hand allows us to look into the first two effects. We define as treated all the children that were in utero in a locust affected area between July 2004 and October 2005. This means that children born between July 2004 and June 2006 were the ones potentially treated if we consider the 9 months in utero.

[6] Lean season is period of seasonal poverty in rural areas between planting and harvest when stocks of food are the lowest, there is no income and families often miss meals.

[7] Locust swarms can also affect stock farmers through a decrease in available pasture.

[8] We found no evidence suggesting that subsequent harvest suffered from locust swarm. The asset depletion channel effect is likely to affect a marginal share of households but this effect could be quite strong on those households.

Figure 2.6: Seasonal Calendar



2.5 Empirical Strategy

The impact of locust plague invasion on child health is investigated using a Difference-in-Differences identification strategy. Treatment and control groups are defined at enumeration area level (household cluster) and all of the variation in childhood exposure to treatment across enumeration areas and birth cohorts is exploited. We first estimate the following equation using ordinary least squares.

$$y_{i(h,t,e)} = \gamma Z_{i(t,e)} + \mu_e + \beta_t + \theta_{hh} X_h + \theta_1 X^{(1)}_{i(h,t,e)} + \theta_2 X^{(2)}_{i(t,e)} + \epsilon_{i(h,t,e)}, \tag{2.1}$$

where i denotes an individual child of cohort (year-month) t, born in household h who lives in enumeration area e. $y_{i(h,t,e)}$ is the height-for-age Z-score of the child i. $Z_{i(t,e)}$ is a dummy for in utero exposure to locust plague invasion for child i. μ_e and β_t are enumeration area and birth cohort fixed effects. $X^{(1)}_{i(h,t,e)}$ and $X^{(2)}_{i(t,e)}$ are observable characteristics at, individual level (birth order, age gap with direct older and younger siblings, etc.) and enumeration area level (Standardized Precipitation-Evapotranspiration Index (SPEI) drought index), respectively. X_h are household level controls (gender and age of household head, wealth index of household, education of the mother, etc.). Standard errors are clustered at enumeration area level.

In the most basic version of the specification shown in equation 2.1, we control for enumeration area fixed effects to account for permanent unobserved characteristics of the place of residence and cohort fixed effects to account for cohort specific shocks. The full version of the main specifications adds controls for the relevant household and child characteristics.

The coefficient γ measures the average difference in changes in height-for-age Z-scores of born children, between locust infested areas and non-infested areas, holding constant all the other relevant characteristics. The implicit assumption behind the identification strategy is that after controlling for cohort fixed effects, enumeration area and household characteristics (or mother fixed effects), and other relevant exogenous covariates, changes in height-for-age Z-scores would be similar across locust infested areas and non-infested areas in absence of the plague.

Interpretation of Coefficient

The coefficient γ does not represent the national impact of locust plague but the average effect with respect to local and cohort averages.

2.6 Results

Table 2.1 shows the estimated effect of being exposed to the adverse effects of locust invasion in utero on height-for-age Z-scores in Mali. Column (1) shows the estimated coefficients for the basic specification from equation 2.1 in which we control only for location fixed effects and cohort fixed effects. The variable of interest is a dummy equal 1 for children born in locust affected areas between July 2004 (beginning of locust plague) and June 2006 (last cohort of children that relied on the 2004-2005 agricultural campaign harvest in utero). The estimated γ coefficient is negative, significant. This coefficient is stable and remains significant after progressively adding controls for child characteristics in column (2) and family characteristics in column (3). The main specification in column (3) suggests that exposed children have, on average, a Z-score 0.24 points below that of non-exposed children. This represents an 18% decrease in the average Z-score for children in our sample and a 20% increase in the average stunting rate.[9]

Column (4) of Table 2.1 controls for region specific time trends and the estimated γ remains stable and significant. Column (5) restricts the estimation sample to children measured during the 2006 DHS survey, hence using older cohorts born before 2004 as non-treated cohorts to build the counterfactual trends. Column (6) does the opposite by restricting the estimation sample to children born after 2004 in which case non-treated cohorts are younger cohorts born after June 2006. The results remain stable in both specifications.

Table 2.4 is the equivalent of Table 2.1 for Senegal. The qualitative results are pretty

[9]Stunting rate is defined as the share of children with height-for-age Z-score smaller than -2 standard deviations. Average stunting rate in our sample is 33%.

similar. The main specification in column (3) suggests that exposed children in Senegal have, on average, a Z-score 0.48 points below that of non-exposed children.

Table 2.1: Impact of Locust Plague on Child Health in Mali

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HAZ	HAZ	HAZ	HAZ	HAZ	HAZ
In utero treatment	-0.216***	-0.180**	-0.239***	-0.245***	-0.298***	-0.221***
	(0.079)	(0.079)	(0.070)	(0.076)	(0.088)	(0.073)
Observations	14,245	14,245	14,225	14,160	9,194	9,889
R-squared	0.161	0.181	0.197	0.212	0.237	0.230
Cohort FE	YES	YES	YES	YES	YES	YES
Enumeration Area FE	YES	YES	YES	YES	YES	YES
Child characteristics	NO	YES	YES	YES	YES	YES
Family characteristics	NO	NO	YES	YES	YES	YES

Full set of controls includes mother's height, education, gender and age of household head, household wealth index, SPEI index, birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.2: Impact of Locust Plague on Child Health in Senegal

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	HAZ	HAZ	HAZ	HAZ	HAZ	HAZ
In utero treatment	-0.396***	-0.436***	-0.477***	-0.457***	-0.368**	-0.500***
	(0.109)	(0.120)	(0.113)	(0.115)	(0.149)	(0.131)
Observations	5,945	5,883	5,785	5,785	2,515	4,023
R-squared	0.132	0.152	0.189	0.197	0.347	0.203
Cohort FE	YES	YES	YES	YES	YES	YES
Enumeration Area FE	YES	YES	YES	YES	YES	YES
Child characteristics	NO	YES	YES	YES	YES	YES
Family characteristics	NO	NO	YES	YES	YES	YES

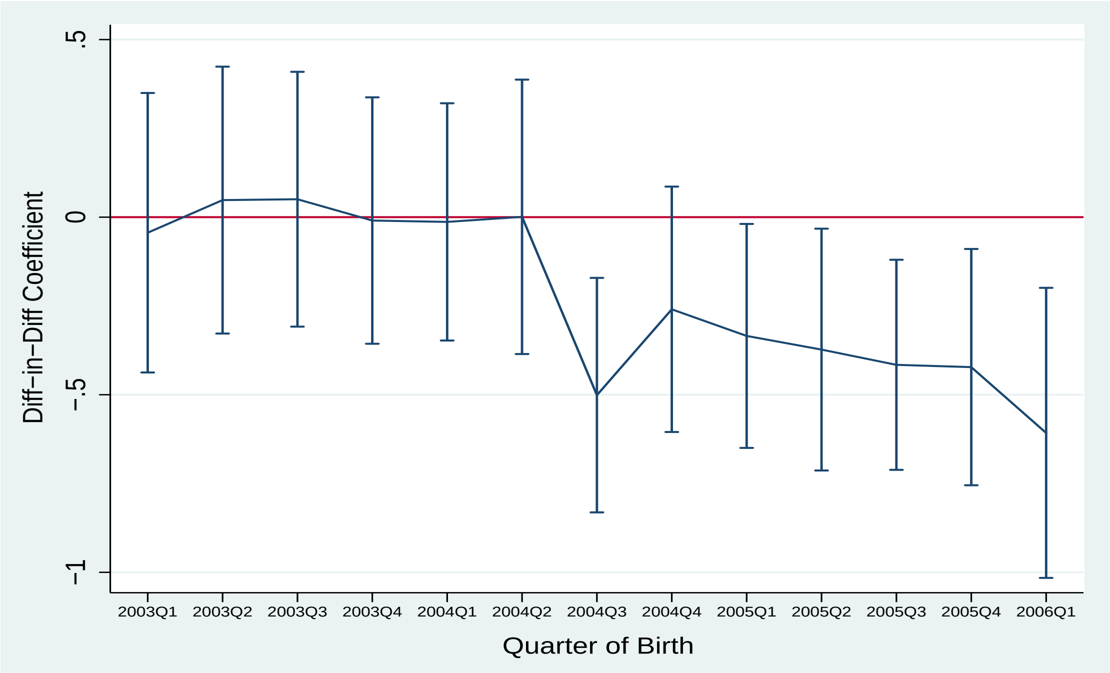
Full set of controls includes mother's height, education, gender and age of household head, household wealth index, SPEI index, birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

Timing of Locust Impact and Placebo Test

Figure 2.7 shows estimated impact of being born in locust infested area by quarter of birth between 2003 and 2006.[10] It shows that children born before the locust invasion started in July 2004, have comparable height-for-age Z-scores between treatment and control groups. This confirms the hypothesis that the estimated impact of exposure to locust plagues is not capturing pre-existing trend differences between treated and non-treated areas. The estimated coefficient is negative and significant for children born in July, August, or September 2004 (first months of locust invasion in Mali). Those children were still relying on the previous harvest that has not been affected by locust. They were, however, potentially affected by speculations on the current agricultural campaign leading to an increase in prices of local food products. The significant effects found for these children support this mechanism. Children born between September 2004 and March 2006 where relying on harvest of the 2004-2005 campaign. They were therefore potentially subject to the supply/income shock due to failure of agricultural production on top of the speculative effects. The negative and significant impact found for these

[10]The 2006 DHS survey data was collected between May and December so we restrict the analysis to children born before April 2006 to prevent survey timing biases from contaminating the estimated effect.

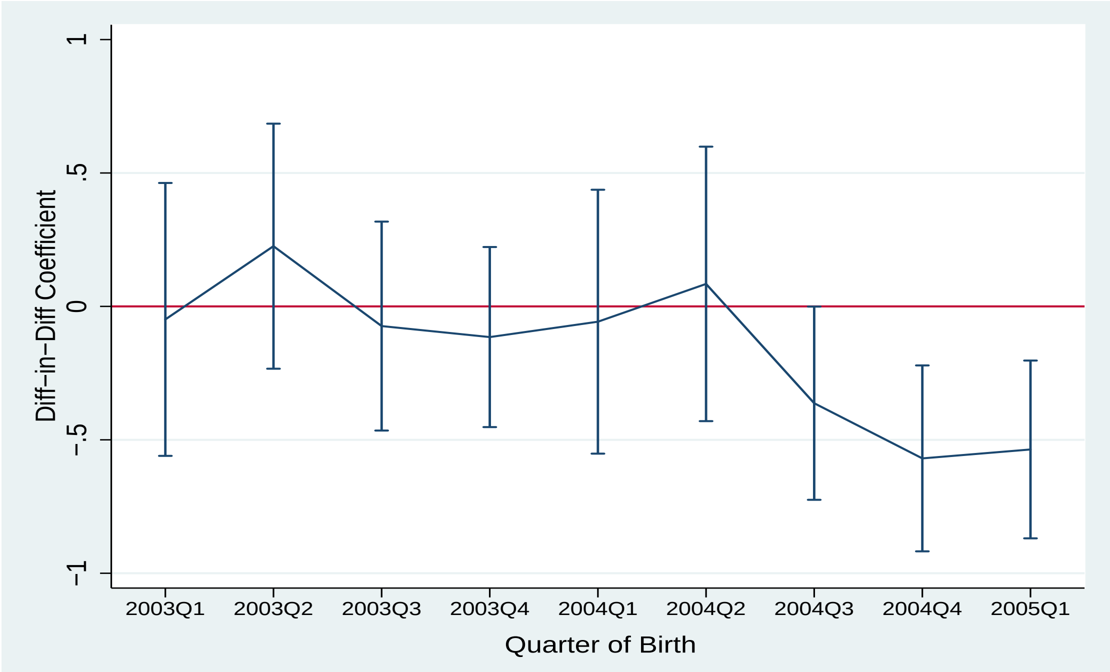
Figure 2.7: Estimated Impact of 2004 Locust Invasion in Mali



cohorts confirm this idea.

Figure 2.8 is the equivalent of Figure 2.7 for Senegal. It also shows there is no pre-existing trend between treated and non-treated areas in Senegal since the estimated impact is small and not significant for cohorts of children born before July 2004. Children born during the third quarter of 2004 in treated areas suffer significant health setbacks due to speculative price effects. Children born after the 2004 harvest also suffer major health setback in treated areas.

Figure 2.8: Estimated Impact of 2004 Locust Invasion in Senegal



Heteregenous Effects and Robustness

Table 2.3: Robustness: Impact of Locust Plague on Child Health in Mali

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Gender		Location		Wealth Index			Years of Residence	
VARIABLES	Male	Female	Rural	Urban	Poor	Middle	Rich	2+ years	Always
In utero treatment	-0.264***	-0.218*	-0.245***	-0.101	-0.225*	-0.339**	-0.163	-0.224***	-0.284**
	(0.087)	(0.112)	(0.081)	(0.143)	(0.121)	(0.159)	(0.167)	(0.070)	(0.108)
Observations	7,193	6,967	10,225	3,935	5,893	2,956	5,474	9,670	5,464
R-squared	0.210	0.231	0.218	0.178	0.229	0.259	0.186	0.215	0.216
Cohort FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Enumeration Area FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Child characteristics	YES	YES	YES	YES	YES	YES	YES	YES	YES
Family characteristics	YES	YES	YES	YES	YES	YES	YES	YES	YES

Dependent variable is HAZ. Full set of controls includes mother's height, education, gender and age of household head, household wealth index, SPEI index, birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

2.7 Conclusion

In this chapter, we show that exposure to locust plagues in utero can lead to substantial health setbacks early in life. Evidence from the 2004 plague in Mali and Senegal suggests

Table 2.4: Robustness: Impact of Locust Plague on Child Health in Senegal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Gender		Location		Wealth Index			Years of Residence	
VARIABLES	Male	Female	Rural	Urban	Poor	Middle	Rich	2+ years	Always
In utero treatment	-0.410**	-0.519***	-0.548***	-0.467**	-0.741***	-0.197	0.061	-0.409***	-0.486**
	(0.177)	(0.159)	(0.147)	(0.228)	(0.155)	(0.311)	(0.233)	(0.120)	(0.202)
Observations	3,065	2,720	3,951	1,832	3,060	1,311	1,442	2,418	1,089
R-squared	0.202	0.223	0.200	0.222	0.232	0.257	0.190	0.214	0.285
Cohort FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Enumeration Area FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Child characteristics	YES	YES	YES	YES	YES	YES	YES	YES	YES
Family characteristics	YES	YES	YES	YES	YES	YES	YES	YES	YES

Dependent variable is HAZ. Full set of controls includes mother's height, education, gender and age of household head, household wealth index, SPEI index, birth order, time gap between conception and the previous and following pregnancies. Robust standard errors in parentheses are clustered at enumeration area level. *** p<0.01, ** p<0.05, * p<0.1.

that there was a strong speculative price effect of locust invasion that kicked in during the plague itself and a local crop supply shock that lasted at least until the next harvest. We find that exposed children have, on average, a Z-score 0.25 points and 0.48 points lower than non-exposed children in Mali and Senegal respectively. The estimated effects are stronger for children born in poor households and more isolated areas.

From policy perspective, this chapter suggest that it is important to address expectations during pest invasions that affect agricultural production in order to prevent speculations and inflation. It is also crucial to provide safety nets to the most vulnerable households that are affected by these shocks to dampen the consequences of these shocks on young children. With alarming climate change prospects, one can expect more frequent extreme weather conditions in the breeding areas of locust leading potentially to more plagues in future. Preparing to limit the consequences of these locust invasions is necessary for policy makers that operate in the Sahel region.

Chapter 3

Access to Power, Political Institutions and Ethnic Favoritism

written jointly with Hannes Mueller.

3.1 Introduction

A large literature in development economics has analysed the role played by ethnic politics for economic development.[1] A central part of the argument has been that economic and political conflict often organises around ethnic lines[2] and there is now ample empirical evidence that this is indeed the case.[3] A part of this literature explains conflict by the fact that ethnic groups regard the nation state as a resource that can be captured and through which resources can be distributed. Evidence for ethnic favoritism, the distribution of state resources towards coethnics, is therefore a possible key ingredient to understand how ethnic politics hinders economic development.

In this chapter we contribute to this literature by showing that ethnic groups with access to executive power seem to benefit economically. In addition, we show that this

[1]See, for example, Easterly and Levine (1997), Alesina et al. (1999) and Alesina et al. (2015). For a review see Alesina and Ferrara (2005).

[2]Fearon and Laitin (2000), Esteban and Ray (2008), Caselli and Coleman (2013)

[3]Montalvo and Reynal-Querol (2005),Esteban et al. (2012), Michalopoulos and Papaioannou (2016)

effect is muted when strong institutional constraints are present at the country level. More specifically, we use data on 564 ethnic groups in 130 countries in the period 1992-2010 to show that night light intensity per capita increases systematically when an ethnic group gains access to executive power and that this effect is dampened by institutional constraints on the executive. A special feature of the data we use is that it does not only measure access to executive power, the extensive margin, but also the intensity to which an ethnic group concentrates executive power, the intensive margin. We use this feature of the data to show that holding more concentrated executive power implies significantly higher night light emissions. When a group holds a monopoly over executive power, for example, it becomes significantly better off than if it shares political power as a senior partner. Similarly, if a group is a junior partner in government it loses less light relative to the most powerful ethnic group than a group which is excluded from government. Again, all these effects at the intensive margin go away when institutional constraints on executive power are in place. Our results are robust to several definitions of institutional constraints and different sample restrictions.

Ethnic favoritism has recently received close attention. Burgess et al. (2015) explore the link between access to power, political institutions and inequality in the allocation of public investments. They provide evidence for ethnic favoritism using variation in political leadership and data on road building in Kenyan districts across the 1963-2011 period. In addition, they show that this favoritism in road investments vanishes during periods of democracy. The authors argue that, in the African context where presidential power is based on ethnicity, even weak democratic institutions translate into a decrease in favoritism towards groups in power as political leaders are forced to share public goods across the wider population. Hodler and Raschky (2014) use data on the birthplace of political leaders and night-light data from 38,427 sub-national regions from 126 countries to provide evidence for ethnic favoritism. They also show that this favoritism is most prevalent in countries with weak political institutions and poorly educated citizens. We follow both papers in our identification strategy by using panel regressions with group fixed effects and country-year fixed effects. However, we also expand the existing findings in two directions. Firstly, we show that favoritism exists at the level of ethnic groups for a large number of countries, i.e. we generalize the findings in Burgess et al. (2015). The larger sample of countries allows us to explore robustness with respect to the measure of political instituions. Secondly, we use different changes in executive power to identify its effect on the economy and the mediating role played by political institutions. Whereas both Hodler and Raschky (2014) and Burgess et al. (2015) focus on the identity of the political leader, we look at different degrees of control over the executive.[4] In our dataset we observe not only whether a group is in power but also how concentrated this power is. In addition, we can see whether a group which is not in power has at least some access to the executive, is completely powerless or is even actively prevented from taking power.

We believe that these results are particularly important in the context of a literature which tries to explain the outbreak of violent ethnic conflict. Esteban and Ray (2008) take the position that "prize-grabbing" on a large scale is frequently at the heart of ethnic conflict. Caselli and Coleman (2013) argue that ethnic markers help enforce group membership which facilitates the organisation of conflict effort. Goldstone et al. (2010) show that political discrimination of ethnic groups is a key variable for predicting the onset of political instability. Cederman et al. (2013) provide the data we use to argue that disadvantaged and advantaged ethnic groups have a higher propensity of entering into conflict but only when there is also inequality in access to power. They also show that a recent loss of power or outright discrimination increases the risk of conflict even further.[5] In this context, favoritism is part of a contest for resources and power which is fought between different groups inside a country. Most recently, similar data has been used by Michalopoulos and Papaioannou (2016) to show that groups which are split by a national boundary are much more likely to be politically discriminated by the central state. They also argue that political discrimination could form part of the link between partinioned groups and violence.

A possible mediating factor of this contest for resources and power are political institutions. Rodrik (1999) argues, for example, that economic shocks trigger conflict in countries which do not have strong political institutions which can mediate an intensified

[4]This also sets us apart from a recent working paper by De Luca et al. (2015).

[5]Buhaug et al. (2014) test the scale dependence of these results. The authors show that countries with very poor (compared to the national average) ethnic groups and those with large discriminated groups from national politics are more likely to experience an armed conflict.

competition for resources. Collier and Hoeffler (2000) argue that fractionalization has negative effects on growth and productivity only in nondemocratic regimes. Acemoglu and Robinson (2005) argue that the adoption of democratic institutions can prevent violent unrest. Besley and Persson (2011) present a theoretical framework in which fighting for control of the state is an important reason for entering in conflict. They argue that conflict is particularly likely if there are no political constraints on how resources can be redistributed. Our study delivers some support for the mechanism suggested by their theoretical framework. We find that even extreme concentrations of executive power do not translate into an increase in economic inequality when executive constraints are present.[6]

The remainder of this chapter is organized as follows. Section 3.2 presents the data and methodology. Section 3.3, presents our main findings and robustness checks. Section Section 3.4 concludes.

3.2 Data and Methodology

The dataset we use is the unified platform for geographical research on war (GROW^{up}) (Girardin et al., 2015) which merges and updates data on Ethnic Power Relations (EPR) from Cederman et al. (2010) with data on night light emissions (NOAA-NGDC, 2013).[7] We use night light per capita intensity as a proxy for economic activity at the ethnic group level.[8] Night light data has the benefit of being available on a yearly basis and of being measured at the local level where there is poor availability of statistical data. In order to calculate light emissions per capita we interpolated population data between 10-year periods at the ethnic group level.

The GROW^{up} dataset captures access to power through the participation of members of relevant ethnic groups in the executive (representation in the presidency, cabinet, and

[6]A complementary interpretation of these findings is that executive constraints (often strong parliaments) work as a deliberation mechanism which prevents the escalation of conflict. This sort of mechanism finds some support in Blattman et al. (2014).

[7]The dataset covers all countries except failed states, overseas colonies and countries with less than 500,000 inhabitants. All politically relevant ethnic groups are included in the dataset. An ethnic group is classified as relevant if at least one political organization claims to represent it in national politics or if its members are subjected to state-led political discrimination.

[8]In this we follow Michalopoulos and Papaioannou (2013).

senior posts in the administration, including the army) and codes it in seven subcategories: discriminated, powerless, self-excluded, junior partner, senior partner, dominant and monopoly.[9] These categories capture how well the group is represented in the executive. For example, if a group is coded as having a *monopoly*, elite members from this group hold monopoly power in the executive to the exclusion of members of other ethnic groups. If the group is a *junior partner*, then representatives of that group share access to executive power with a more powerful group (the *senior partner*).

In Table 3.1 we show the example of Kenya. The table shows the seven point scale coding for the seven relevant Kenyan ethnic groups from 1963 to 2013 (which includes the period studied by Burgess et al. (2015)). For simplicity, we have aggregated the yearly data into 6 episodes with constant values. The power status reaches from 1 (discrimated) to 5 (senior partner) and there is both a lot of heterogeneity between groups and across time within the same ethnic group. Transitions from and to democracy took place in 1969 and 1992, i.e. these changes are not reflected by changes in the access to executive power as captured by the EPR data. For comparison we also report the ethnicity of the Kenyan political leader as a blue shade. This is the variation used by Burgess et al. (2015) and Hodler and Raschky (2014).

Table 3.1: Ethnic groups power access in Kenya

Ethnic groups	1963-1966	1967-1978	1979-2002	2003-2005	2006-2007	2008-2013
Kalenjin-Masai- Turkana-Samburu	4	4	5	4	4	4
Kamba	4	4	4	4	4	4
Kikuyu-Meru-Emb	5	5	1	5	5	5
Kisii	4	4	4	2	2	4
Luhya	4	4	4	4	4	4
Luo	4	1	1	4	1	5
Mijikenda	2	2	4	4	4	4

Notes: This table displays variable "status pwrrank" in GROW^{up} dataset for relevant ethnic groups in Kenya from 1963 to 2013. It gives the political status of each ethnic group, ranked on a 7 point scale: 1 (Discriminated), 2 (Powerless), 3 (Self-excluded), 4 (Junior Partner) and 5 (Senior Partner) in government. The ethnic group of the leader is marked by the (blue) shade for each period. First president Jomo Kenyatta (ethnic kikuyu) was in power from 1963 to 1978. He was replaced by Daniel arap Moi (ethnic Kalenji) who stayed in power from 1979 to 2002 followed by Mwai Kibaki (ethnic Kikuyu) who stayed in office from 2003 to 2013. Kenya experienced a first transition from autocracy to democracy (in December 1969) under president Kenyatta. It switched back to democracy in December 1992 under president Moi.

[9]For the coding of this variable, the experts focus on the most relevant part of the executive (e.g., in a military dictatorship, power over the army; in presidential systems, the senior cabinet). They look at absolute access to power irrespective of the question of under- or overrepresentation relative to the demographic size of an ethnic category. See the appendix for a more detailed description of the different subcategories.

Our first main question is whether executive power translates into economic favoritism. Our dataset allows us to test this on three dimensions. First, we analyse how the change of groups into and from executive power affects the local economy. For this purpose we code the variable $leader_{i,j,t}$ which takes a value of 1 for the ethnic group j with the maximum executive power in country i and year t. This can be either a group that has a monopoly on executive power, is dominant or is a senior partner in government. Secondly, we can look at changes at the intensive margin because we can split the variable $leader_{i,j,t}$ up into its three subcomponents and add them separately. Thirdly, we can look at the downside of having no access to executive power by using dummies for groups who are junior partners, powerless, discriminated or self-excluded.

Our econometric specification assumes that log night light emissions per capita in country i, group j and year t are given by:

$$log(light\ per\ capita)_{i,j,t} = \beta * leader_{i,j,t} + C_{i,t} + \eta_j + \epsilon_{i,j,t}$$

where $log(light\ per\ capita)_{i,j,t}$ is the logarithm of per capita night light intensity for ethnic group j in country i at time t. The coefficients $C_{i,t}$ capture a set of country-year dummies to control for shocks and changes that are common to all ethnic groups within any given country, as well as for changes in satellites and their sensor settings. Importantly, this implies that the coefficient on $leader_{i,j,t}$ needs to be interpreted in relative terms, i.e. in comparison to the average in the same country and year.[10] We also include ethnic group fixed effects to control for ethnic groups permanent unobserved characteristics.

The mechanism we have in mind builds on the previous findings in Burgess et al. (2015) and Hodler and Raschky (2014). Our empirical specification assumes that gaining power will not affect economic growth in a region but will instead change the relative level of output up and above country-wide changes in economic output. This would be in line with the idea that public spending gets reallocated but is not effective in generating a different growth trend. In other words, we propose that groups which manage to redirect funding to their areas do not use this spending systematically to improve growth but instead to boost (state) consumption.

[10] Our results are also robust to using lagged values of the right-hand-side variables

We identify the effect of political power from changes at the ethnic group level. Table 3.2 displays the full variation of the seven-point scale at this level. Rows display access to power of group j in period t while columns display access to power of the same group in period $t+1$. The number in each cell displays the count of transitions between the different levels of access to power in our data. There were, for example, 3616 transitions from "powerless" to "powerless", 46 transitions from "powerless" to "junior" and 25 transitions from "junior" to "powerless". Table 3.2 clearly indicates a dominance of transitions within the same category. A group that has no access to power in year t has most likely no access to power in year $t+1$ as well. In addition, large jumps in the access to power are rare. Most transitions are between adjacent categories, i.e. from powerless to junior or from junior to senior. This is important to keep in mind as our identification strategy relies on these transitions.

The last two columns of Table 3.2 provide summaries. The column "transitions out" displays the number of transitions from a power status in t to another power status in $t+1$ while "transitions into" are the number of transitions from another access to power in t to the respective category in $t+1$. For example, in our data we have 32 ethnic groups which escape being discriminated and 19 groups who become discriminated. Overall we have 232 changes in the access to power in over 9000 observations. This is important as it implies that changes in power are rare and important events for which we might expect changes in policies. The downside is that we have few transitions in some categories. The effect of holding a monopoly of power, for example, will be estimated from just 14 transitions.

Table 3.2: Transitions

		access to power in t+1							transitions out	transitions to
		self excluded	discriminated	powerless	junior	**senior**	**dominant**	**monopoly**		
power in t	**monopoly**	0	1	0	1	3	6	493	11	3
	dominant	0	0	1	8	17	647	0	26	20
	senior	0	2	6	18	1134	10	2	38	58
	junior	0	7	25	2120	20	2	1	55	85
	powerless	1	9	3616	46	11	1	0	68	45
	discriminated	1	734	12	11	7	1	0	32	19
	self excluded	211	0	1	1	0	0	0	2	2

Notes: This table displays changes in access to political power for all ethnic groups in our sample. For definitions see the text. The ordering reflects roughly the extent of executive power a group has. Bold indicate categories which we code as "leaders" in the first part of our analysis. Rows display access to power in period t while columns display access to power in period t+1. There were, for example, 3616 transitions from the status "powerless" to the status "powerless", 46 transitions from "powerless" to "junior" and 25 transitions from "junior" to "powerless". The last two columns provide summaries of the table. "transitions out" are the total number of transitions from a power status in t to another power status in t+1. "transitions into" are the total number of transitions from another power status in t to the respective power status in t+1.

We merge this data with data on political institutions from the Polity IV dataset

(Marshall and Gurr, 2013). To capture democratic institutions we add an interaction effect with a dummy for strong executive constraints in country i and year t, $constraints_{i,t}$, to the regression above. Most of the literature on political institutions treats democracy as an aggregate outcome based on the index in Polity IV. We use a more disaggregated approach because, from a theoretical perspective it makes sense to focus on the extent to which executive power is constrained by political institutions. The use of institutions as a constraint on power is an old idea. Mill (1869), for example, described a limit to the power of a ruler that can be achieved through

> *"[...] establishment of constitutional checks, by which the consent of the community, or of a body of some sort, supposed to represent its interests, was made a necessary condition to some of the more important acts of the governing power"*

We define strong executive constraints through the variable *xconst* from Polity IV. This variable captures the extent of institutionalized constraints on the decision-making powers of chief executives. Such limitations may be imposed by any "accountability groups". In Western democracies these are usually legislatures. Other kinds of accountability groups are the ruling party in a one-party state or an independent judiciary. The variable *xconst* is coded from one to seven. However, there is no reason to believe that each change in the index has the same meaning. At a score of $xconst = 7$ the constraint on the executive is clearly biting as *"Accountability groups have effective authority equal to or greater than the executive in most areas of activity." (Polity IV, Coding Manuel)* The idea is that even concentrated executive power cannot be translated into economic favoritism when strong constraints are in place.[11] In our baseline specification, we define strong executive constraints through a value of *xconst* equal to 6 or 7 and run robustness checks with $xconst = 7$. A value of 6 is an intermediate step which adds additional observations to the camp of strong executive constraints without changing the definition substantively.

[11] In this context it is worth stressing that discrimination in our coding above is discrimination in the access to power and not a policy outcome.

However, we will also run robustness checks with other dimensions of political institutions by using the competitiveness of executive recruitment and the openness of executive recruitment from the Polity IV dataset. Table 3.3 provides summary statistics for all the variables used in our analysis including the additional institutional measures.

Table 3.3: Summary Statistics of Main Variables

Variables	Obs	Mean	SD	Min	Max
log light per capita	9210	-4.331	2.019	-19.928	0.972
self excluded	9210	0.023	0.150	0	1
discriminated	9210	0.082	0.275	0	1
powerless	9210	0.398	0.490	0	1
junior partner	9210	0.240	0.427	0	1
senior partner	9210	0.130	0.336	0	1
dominant	9210	0.073	0.259	0	1
monopoly	9210	0.054	0.226	0	1
leader	9210	0.256	0.437	0	1
strong executive constraint	9210	0.374	0.484	0	1
high competitiveness of executive recruitment	9210	0.379	0.485	0	1
high openness of executive recruitment	9210	0.790	0.408	0	1

3.3 Results

Table 3.4 shows our main results. In Column (1) we find that, relative to other ethnic groups in the same country and year, light per capita increases by about 6.6 percent when a group starts to control executive power. If we assume a log relationship between GDP and light as in Henderson et al. (2012) this implies that GDP per capita increases by about 2.2 percent.[12] In Column (2) we add the interaction term $leader_{i,j,t} \times constraints_{i,t}$ and find that the effect of executive power on local economies is significantly lower when strong executive constraints are present. Moreover, the sum of the coefficients reported in column 2 is not statistically different from 0 which implies that there is evidence of ethnic favoritism only in absence of constraints on the executive power.

Next, we decompose the variable $leader_{i,j,t}$ into its three subcomponents (senior partner, dominant and monopoly) and check whether effects of leadership are heterogeneous according to how concentrated executive power is. Column (3) shows that there is a perfect match between the concentration of power and favoritism in terms of per capita

[12] Hodler and Raschky (2014) show that this relationship also holds for subnational regions.

light. Ethnic groups that enjoy a political monopoly on power receive 13.5 percent more night light per capita. Groups with dominant access to power and senior partners in government receive respectively 8.2 and 5.4 percent more light per capita.[13] The more concentrated executive power, the more does the ethnic group benefit economically. Note, that some of this effect could come from changes in light emissions in excluded categories. The definition of a monopoly of executive power, for example, directly implies that other groups will be powerless. We will return to this point below.

In column (4) of the same table, we add interaction effects between the different categories of power and our measure of political constraints. A lot of the variation here is driven by changes in constraints which, as the example of Kenya makes clear, do not generally coincide with changes in executive power. The results we find here are striking. Under weak executive constraints the magnitude of the effect of power amplifies for all three categories. A monopoly over executive power, for instance, increases light per capita by 21.6 percent in country/years with weak executive constraints. If we use the elasticity proposed by Henderson et al. (2012) this suggests that groups with a monopoly of power experience a relative expansion of GDP per capita by about 7 percent compared to the national average. However, we cannot reject the hypothesis that this difference completely vanishes under strong executive constraints, again for all three categories. Under strong executive constraints, there are no differences in night light per capita between the control group and groups with a monopoly, dominant power status, or senior partners.[14] In Column (5) we show that results also do not change if we control for area, population and urbanization time trends.[15]

Our seven point measure of access to political power also allows us to look at the absence of political power: junior partners in government, powerless groups, discriminated groups and groups that are self-excluded. Consistent with the findings in Table 3.4 we find lower night light per capita when looking at these groups. Table 3.5, column (1) shows

[13] The difference between the coefficients on monopoly and senior partner is significant at 1%.

[14] Groups with a monopoly of power can be found in countries with strong executive constraints. We have 19 such countries in our sample: Albania, Argentina, Austria, Bulgaria, Chile, Cyprus, Ecuador, El Salvador, France, Greece, Hungary, Japan, Panama, Paraguay, Poland, Slovakia, Turkey, Uruguay and Venezuela.

[15] Area, population and urbanization time trends are respectively area, population and urbanization at the beginning of the period multiplied by a time trend.

Table 3.4: Executive Power and Ethnic Favoritism

VARIABLES	(1) light per capita	(2) light per capita	(3) light per capita	(4) light per capita	(5) light per capita
leader	0.066***	0.093***			
	(0.018)	(0.022)			
leader * strong executive constraints		-0.070***			
		(0.016)			
senior partner			0.054***	0.064***	0.064***
			(0.018)	(0.024)	(0.024)
dominant			0.082**	0.094**	0.091**
			(0.032)	(0.043)	(0.043)
monopoly			0.136***	0.216***	0.222***
			(0.040)	(0.048)	(0.047)
senior partner * strong executive constraints				-0.037*	-0.035*
				(0.020)	(0.021)
dominant * strong executive constraints				-0.062**	-0.065**
				(0.032)	(0.031)
monopoly* strong executive constraints				-0.265***	-0.273***
				(0.059)	(0.059)
Observations	9,210	9,210	9,210	9,210	9,177
R-squared	0.991	0.991	0.991	0.991	0.991
Country-Year FE	YES	YES	YES	YES	YES
Group FE	YES	YES	YES	YES	YES
(population, area and urbanisation) x trend	NO	NO	NO	NO	YES

Robust standard errors clustered at country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. "light per capita" is the ln of night light emissions per capita. "senior partner", "dominant" and "monopoly" are dummies that capture increasing executive power of the ethnic group. "leader" is a dummy that takes 1 if ethnic group has the highest level of access to power in the country/year ("senior partner", "dominant" or "monopoly"). "strong executive constraints" is a dummy which captures values of xconst>5. Regressions use population weights.

that junior groups receive 4.9 percentage points less light on average than groups in power. This decreases to 12.8 percent less light for discriminated groups.[16] In column (2) we show that, in line with our earlier results, these effects are stronger under weak executive constraints than under strong executive constraints. Under weak executive constraints, discriminated groups receive 14 percent less light per capita than the average. Under strong executive constraints, it makes no difference whether a group is a junior partner in government, powerless or discriminated. In column (3) we show that these results are robust to the inclusion of area, population and urbanization time trends.

A possible concern about the identification of ethnic favoritism in our empirical framework is the potential bias induced by reverse causality or omitted variable bias. Groups might gain power, for example, because they have become economically more powerful.

[16]The coefficients on self-excluded groups are identified of only a handful of transitions and we therefore do not discuss them.

Table 3.5: Executive Power and Ethnic Favoritism - The Downside

VARIABLES	(1) light per capita	(2) light per capita	(3) light per capita
junior partner	-0.049**	-0.071***	-0.071***
	(0.019)	(0.027)	(0.026)
powerless	-0.085***	-0.117***	-0.114***
	(0.030)	(0.037)	(0.037)
discriminated	-0.128***	-0.141***	-0.146***
	(0.035)	(0.036)	(0.036)
self excluded	-0.302***	-0.336***	-0.340***
	(0.056)	(0.060)	(0.060)
junior partner * strong executive contraints		0.055**	0.054**
		(0.023)	(0.024)
powerless * strong executive constraints		0.082***	0.085***
		(0.028)	(0.027)
discriminated strong executive constraints		0.064	0.066
		(0.066)	(0.066)
self excluded * strong executive constraints		0.191***	0.172**
		(0.067)	(0.067)
Observations	9,210	9,210	9,177
R-squared	0.991	0.991	0.991
Country-Year FE	YES	YES	YES
Group FE	YES	YES	YES
(population, area and urbanisation) x trend	NO	NO	YES

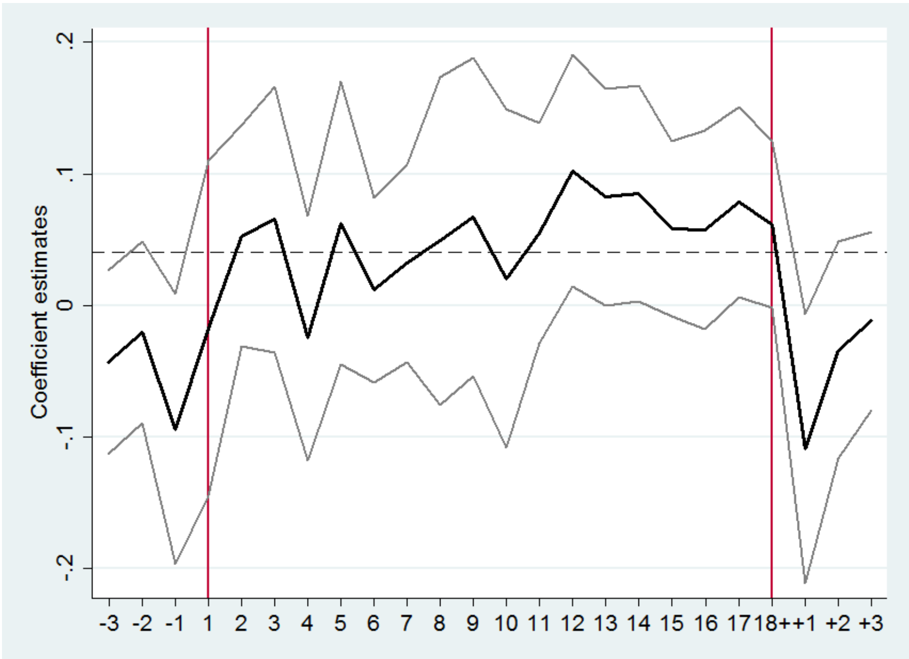
Robust standard errors clustered at country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. "light per capita" is the ln of night light emissions per capita. "junior partner", "powerless", "discriminated" and "self excluded" are dummies that capture decreasing access to notional level executive power of the ethnic group. "strong executive constraints" is a dummy which captures values of xconst>5. Regressions use population weights.

However, relative changes in economic activities across ethnic groups in a given country are more likely to be gradual while most changes in political leadership occur after elections, military coups or natural death of leaders. Also, it is hard to reconcile the absence of a relationship under strong executive constraints with reverse causality.

In any case, if economic changes anticipate political changes we would expect ethnic groups in power to have more intense light per capita already before holding the leadership and, perhaps, a declining trend after losing power. To test whether there is such pattern in the data, we look at ethnic groups that are about to become leaders or have been leaders recently to see whether they look different in these years compared to other years in which they were not leaders. To do so, we decompose the $leader_{i,j,t}$ dummy in column (1) of Table 3.4 into a set of dummies for future, current, and past political leadership . In particular, we add 3 dummies for the last 3 years before groups gain power and 3 dummies for the first 3 years after an ethnic group lost power. We also add 17 dummies for the first 17 years in which an ethnic group is in power, and one dummy for all subsequent years in power.

The results are shown in Figure 3.1 and suggest that there is no difference in night light per capita before or after an ethnic group becomes leader compared to the other years in which it is not in power. After getting access to power, we observe a significant increase in per capita light intensity which takes place within a few years and then remains fairly stable across time in power. These results suggest that there are no pre-trends before gaining power or negative trends after losing power. It is rather that changes in political power coincide with a relatively sharp and seizable change in economic activity.

Figure 3.1: The dynamics of ethnic favoritism



Notes: The graph shows the point estimates of a series of dummy variables accounting for the individual three years before an ethnic group becomes a leader (-3, -2, -1), the first 17 years of being a leader (1-17), the eighteenth and all subsequent years (18+), and the first three years after the end of the group's leadership (+1, +2, +3). The black line plots the point estimates, and the gray lines indicate the upper and lower limits of the 95% confidence interval. These estimates come from a single fixed effects regression, where log light per capita is regressed on the 24 dummy variables and the full set of country-year and ethnic group dummy variables. The vertical lines in red indicate the first and the last dummy variable representing leadership of ethnic groups. The horizontal dashed line indicates the estimate of leader coefficient in Table 3.6, Column (2).

A simple summary of these results is in columns (1) and (2) of Table 3.6. Here we collapse the dummies and show just the coefficients on the dummies for the time before and after a switch into government, i.e. a switch from less power than a senior partner to more. Column (1) displays the estimated coefficients on the year right before and right after the switch. Column (2) displays the estimated coefficient for a dummy that captures the three years before and after. We find no significant deviation from 0 around the switching date. The same is true if we move the cut-off of what defines leader to include junior partners (columns (3) and (4) in Table 3.6) or if we drop senior partners from the definition (columns (5) and (6)).[17]

Table 3.6: Dynamics of Ethnic Favoritism

	(1)	(2)	(3)	(4)	(5)	(6)
	Senior +		Junior +		Dominant +	
VARIABLES	light per capita	light per capita	light per capita	light per capita	light per capita	light per capita
Leader	0.053***	0.041*	0.075***	0.088***	0.084***	0.077***
	(0.019)	(0.024)	(0.026)	(0.032)	(0.025)	(0.028)
Lost leadership recently	-0.046	-0.033	0.006	-0.002	0.086	0.038
	(0.043)	(0.032)	(0.066)	(0.052)	(0.067)	(0.030)
About to gain leadership	-0.056	-0.048	0.004	0.031	0.059	0.019
	(0.051)	(0.033)	(0.045)	(0.031)	(0.057)	(0.045)
Observations	9,186	9,086	9,195	9,120	9,180	9,068
R-squared	0.991	0.991	0.991	0.991	0.991	0.991
Country-Year FE	YES	YES	YES	YES	YES	YES
Group FE	YES	YES	YES	YES	YES	YES

Robust standard errors clustered at country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. "light per capita" is the ln of night light emissions per capita. "leader" is a dummy that takes 1 if ethnic group has access to power above a certain threshold. In column (1) and (2) leader is defined as senior partner or above ("senior partner", "dominant" or "monopoly") and is called "senior+". In columns (3) and (4) we lower this threshold by one level to include Junior Partners ("Junior +") . In columns (5) and (6) we restrict leader status to groups that are dominant or have monopoly of power ("Dominant+"). "lost leadership recently" is a dummy that takes value 1 if an ethnic group lost leadership status the previous year (columns (1), (3) and (5)) or within the last 3 years (columns (2), (4) and (6)). "about to gain leadership" is a dummy that takes value 1 if an ethnic group will be in power next year (columns (1), (3) and (5)) or within the next 3 years (columns (2), (4) and (6)). Regressions use population weights.

Robustness Checks In Table 3.7 we produce a number of robustness checks. The first 3 columns show robustness to different institutions. In column (1) we use the cut-off of 7 for our definition of strong executive constraints and our results are broadly consistent with our main results.[18] In column (2) we use the level of competitiveness of executive recruitment to define strong political institutions. Political constraints are assumed to be present if the variable *xrcomp* in Polity IV is equal to its highest value of 3; meaning that chief executives are typically chosen in competitive elections matching two or more major

[17]Most of the coefficients are positive now but there is no evidence for a trend around adoption.

[18]This is closely in line with theoretical and empirical work at the country level (Besley and Persson (2011)) who also use the cut-off at 7.

parties or candidates. Results are, again, similar to our main findings. In column (3), we perform a similar exercise with openness of executive recruitment as institutional feature and our results are not robust.[19] This suggests that what prevents ethnic favoritism is the accountability of the executive power due to institutional constraints or incentives for re-election rather than the possibility for anybody to be part of the executive.[20]

We also show robustness to sample restrictions in columns (4) to (7) of Table 3.7 to make sure our findings are not driven by a specific type of country. A problem with looking at different samples is that we need to make sure that enough transitions remain in the sample. In order to do this we drop one income quartile at a time.[21] In column (4) we drop the poorest countries, in column (5) we add these countries back and only drop the second-poorest countries and so on. Standard errors and the size of the coefficient varies but in all columns we get a positive and rising coefficient with increasing power in weak executive constraints and a negative and falling coefficient on the interaction term with strong executive constraints.

In addition, we looked at a more flexible functional form in terms of poulation by having log light as a dependent variable and adding log population on the right hand side. Results are reported in Appendix Table C1. All our main results are robust to this specification. We also implemented additional robustness checks by adding the lagged value of light per capita as a control and discarding population weights. Our main results are also robust to this.

[19]Openness in Polity IV means that the "recruitment of the chief executive is open to the extent that all the politically active population has an opportunity, in principle, to attain the position through a regularized process". We define strong political institution here as a dummy that takes 1 if variable "xropen" in Polity IV data is equal to 4 meaning that chief executives are chosen by elite designation, competitive election, or transitional arrangements between designation and election.

[20]Our results are also not robust to using the aggregate polity score from Polity IV dataset as an indicator for strong institutions.This result is in line with findings in De Luca et al. (2015) and suggests that the level of constraints on the executive is the main institutional feature that mitigates ethnic favoritism.

[21]We use average GDP per capita in 1990 and 2000 to build these quartiles. These are the only two years in GROW^{up} dataset for which we have both data on GDP and population.

Table 3.7: Robustness to measures of good institutions and sample restrictions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Political Institutions			Sample restricted by dropping the following quartiles:			
				Q1	Q2	Q3	Q4
VARIABLES	light per capita	light per capita	light per capita	light per capita	light per capita	light per capita	light per capita
senior partner	0.052**	0.055**	0.075**	0.100***	0.037	0.077***	0.063***
	(0.022)	(0.025)	(0.035)	(0.032)	(0.028)	(0.028)	(0.024)
dominant	0.089**	0.090**	0.053	0.164***	0.057	0.107**	0.090**
	(0.035)	(0.039)	(0.074)	(0.029)	(0.061)	(0.047)	(0.044)
monopoly	0.179***	0.177***	0.145*	0.229***	0.246***	0.196***	0.215***
	(0.047)	(0.051)	(0.081)	(0.042)	(0.075)	(0.062)	(0.048)
senior partner * strong institutions	-0.003	-0.011	-0.025	-0.017	-0.052**	-0.039	-0.036*
	(0.023)	(0.026)	(0.031)	(0.025)	(0.023)	(0.025)	(0.020)
dominant * strong institutions	-0.058*	-0.048*	0.039	-0.088***	-0.028	-0.088***	-0.052
	(0.033)	(0.027)	(0.069)	(0.022)	(0.048)	(0.033)	(0.034)
monopoly strong institutions	-0.178***	-0.108*	-0.007	-0.234***	-0.250***	-0.385***	-0.276***
	(0.063)	(0.059)	(0.083)	(0.059)	(0.090)	(0.099)	(0.061)
Observations	9,210	9,210	9,210	6,641	6,232	7,034	7,723
R-squared	0.991	0.991	0.991	0.990	0.993	0.990	0.987
Country-Year FE	YES	YES	YES	YES	YES	YES	YES
Group FE	YES	YES	YES	YES	YES	YES	YES

Robust standard errors clustered at country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. First 3 columns show robustness to different definitions of good political institutions and the last ones show robustness to sample restrictions. "light per capita" is the ln of night light emissions per capita. "senior partner", "dominant" and "monopoly" are dummies that capture increasing executive power of the ethnic group. "leader" is a dummy that takes 1 if ethnic group has the highest level of access to power in the country/year ("senior partner", "dominant" or "monopoly"). "strong political institutions" is a dummy which captures values of xconst=7 in column (1), xrcomp=3 in column (2) and xropen=4 in column (3). In columns (4) to (7) we use our main definition of good political institutions i.e. xconst>5. Column (4) drops from our sample ethnic groups in countries that belong to the lowest GDP per capita quartile Q1. Column (5) drops countries in the lower middle income quartile Q2 while column (6) drops those in the upper middle quartile Q3. In column (7) we drop the ethnic groups in countries that belong to the top income quartile Q4. Regressions use population weights.

3.4 Conclusions

In this chapter we have shown that ethnic groups which hold executive power at the country level generate more light per capita locally and that this effect is dampened or vanishes in environments with strong institutional constraints on the executive.

Our data allows us to study higher and lower concentrations of executive power and we find that groups that concentrate political power in the executive also benefit economically. Again, this effect is dampened or vanishes in countries with strong executive constraints. This lends additional credibility to existing findings in the literature which suggest a causal link from political representation to economic benefits and highlight the role played by checks and balances in restricting favoritism.

Our estimates are most precise when defining institutional constraints through high scores on the dimension of executive constraints. This captures, for example, that a legislature, ruling party, or council of nobles initiates much or most important legislation. Surprisingly, these sort of institutional constraints seems to constrain favoritism even if the measured concentration of executive power grows. This strengthens the view that political institutions are an important mediating factor of internal conflicts.

The existing data does not, however, allow us to make sharp distinctions regarding which institutions exactly restrict favoritism or how they do it. Further research on the mechanism by which these institutions prevent favoritism appears to be a logical next step.

Appendix: Chapter 1

A.1 Importance of the Smoothing Parameter in KDE

To illustrate the importance of the smoothing parameter h in KDE, I draw 500 data points from a bi-modal distribution given by a mixture of 2 normal distributions with means 5 and 15 and standard deviation 3. The underlying distribution is shown in panel (a) of Figure A1. Panel (b) shows the estimated density with a small and large smoothing parameter h. The first graph still shows some spurious variations from the data while the second one over-smooths the distribution to the extent of almost not reflecting its bi-modal nature.

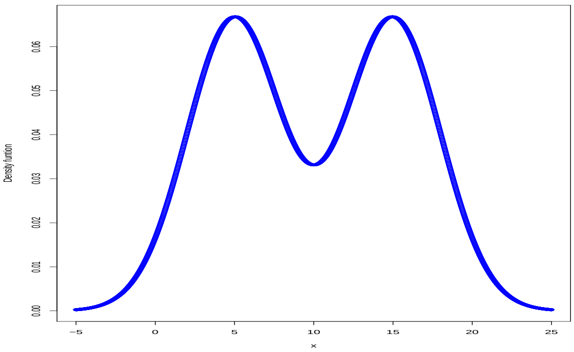
A.2 Bootstrap

The idea of using bootstrap resampling to choose a smoothing bandwidth has been introduced by (Taylor, 1989). The bootstrap approach is used in conjunction with the $MISE(h)$ as a target criterion. The basic idea is to construct a "reference" density estimate of the data at hand, repeatedly simulate data from that reference density, and calculate the empirical integrated squared error at each iteration; doing so at different bandwidths. The bandwidth that minimises the bootstrap-estimated MISE is taken as the optimal value.

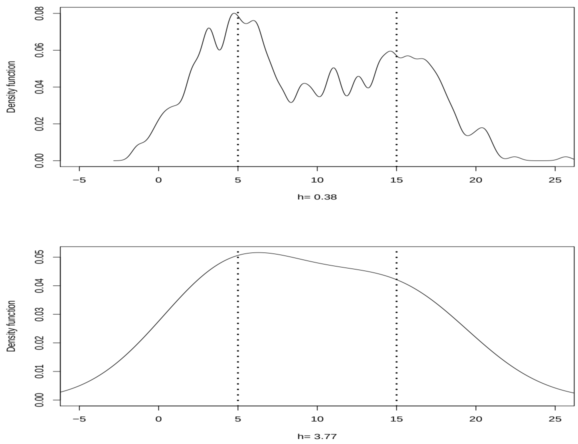
- Select a pilot bandwidth g and compute estimator $\hat{f}_g$ of f

- Draw bootstrap samples X_1^*, ..., X_J^* from $\hat{f}_g$

Figure A1: The role of the smoothing parameter



(a) Underlying density



(b) KDE with small and large bandwidth

- Compute the bootstrap version of the $MISE$ and minimise it over h:

$$J^{-1}\sum_{j=1}^{J}\int[\hat{f}_h(y|X_j^*)-f_g(y|X)]^2dy$$

- Set new pilot bandwidth to the value h_0 that minimizes the $MISE$ and iterate untill it converges

A.3 Descriptive Statistics

Table A1: Household characteristics in Ivory Coast by enumeration area (EA).

	URBAN			RURAL		
	Non war EA (1)	War EA (2)	Difference (1) - (2)	Non war EA (1)	War EA (2)	Difference (1) - (2)
Number of children born in household	3.041	2.813	0.228	3.568	3.480	0.088
	(0.208)	(0.078)	(0.218)	(0.127)	(0.078)	(0.148)
Wealth index	3.859	4.181	-0.322***	2.267	2.059	0.208*
	(0.129)	(0.048)	(0.135)	(0.097)	(0.053)	(0.110)
Mother's height (centimeters)	158.334	159.775	-1.441***	158.097	158.413	-0.316
	(0.656)	(0.243)	(0.686)	(0.337)	(0.318)	(0.461)
Mother's years of education	2.023	2.990	-0.967***	0.774	1.381	-0.607***
	(0.396)	(0.182)	(0.428)	(0.136)	(0.116)	(0.178)
Age of household head	44.050	45.872	-1.822	45.436	46.156	-0.720
	(2.075)	(0.649)	(2.131)	(0.700)	(0.599)	(0.917)
Female household head	0.186	0.204	-0.017	0.080	0.150	-0.070***
	(0.040)	(0.018)	(0.043)	(0.016)	(0.017)	(0.023)
Number of observations	220	1385	1605	979	2379	3358

War EA are enumeration areas with at least 1 conflict event within a 50 km radius between 1997 and 2012. Robust standard errors in parentheses, clustered at the enumeration area (EA) level. * significant at 10%, ** significant at 5%, and *** significant at 1%.

Table A2: Household characteristics in Ivory Coast by treatment status in war affected enumeration areas (EA)

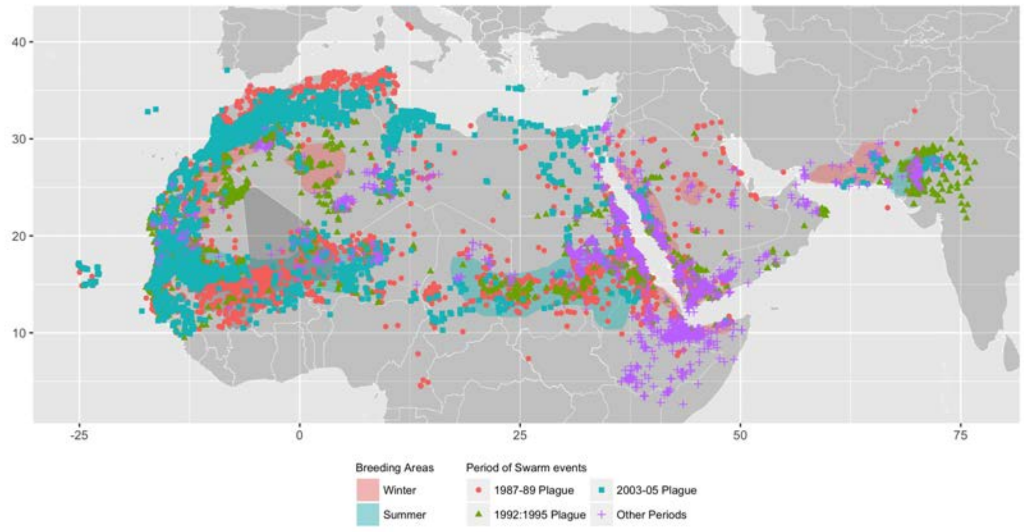
	URBAN			RURAL		
	Never treated HH (1)	Treated HH (2)	Difference (1) - (2)	Never treated HH (1)	Treated HH (2)	Difference (1) - (2)
Number of children born in household	2.752	2.832	-0.080	3.296	3.590	-0.293***
	(0.133)	(0.093)	(0.162)	(0.102)	(0.097)	(0.129)
Wealth index	4.224	4.168	0.056	2.047	2.066	-0.018
	(0.060)	(0.057)	(0.074)	(0.070)	(0.063)	(0.079)
Mother's height in centimeters	160.064	159.684	0.380	159.012	158.056	0.956
	(0.439)	(0.269)	(0.481)	(0.388)	(0.439)	(0.578)
Mother's years of education	3.224	2.917	0.307	1.323	1.416	-0.093
	(0.320)	(0.200)	(0.347)	(0.163)	(0.136)	(0.186)
Age of household head	45.245	46.071	-0.826	45.320	46.653	-1.332
	(1.465)	(0.682)	(1.564)	(0.718)	(0.804)	(1.022)
Female household head	0.233	0.194	0.038	0.142	0.155	-0.013
	(0.029)	(0.022)	(0.036)	(0.021)	(0.021)	(0.027)
Number of observations	331	1054	1385	888	1491	2379

War EA are enumeration areas with at least 1 conflict event within a 50 km radius between 1997 and 2012. Treated HH (households) are families who had at least 1 child exposed (in utero or during first year of life) to the violence when it happened in their area. Non treated families had all their children before conflict started or after conflict started but there was no violent event in their area of residence around the birth of all their children. Robust standard errors in parentheses, clustered at the enumeration area (EA) level. * significant at 10%, ** significant at 5%, and *** significant at 1%.

Figure A2: Partition of Ivory Coast during Conflict

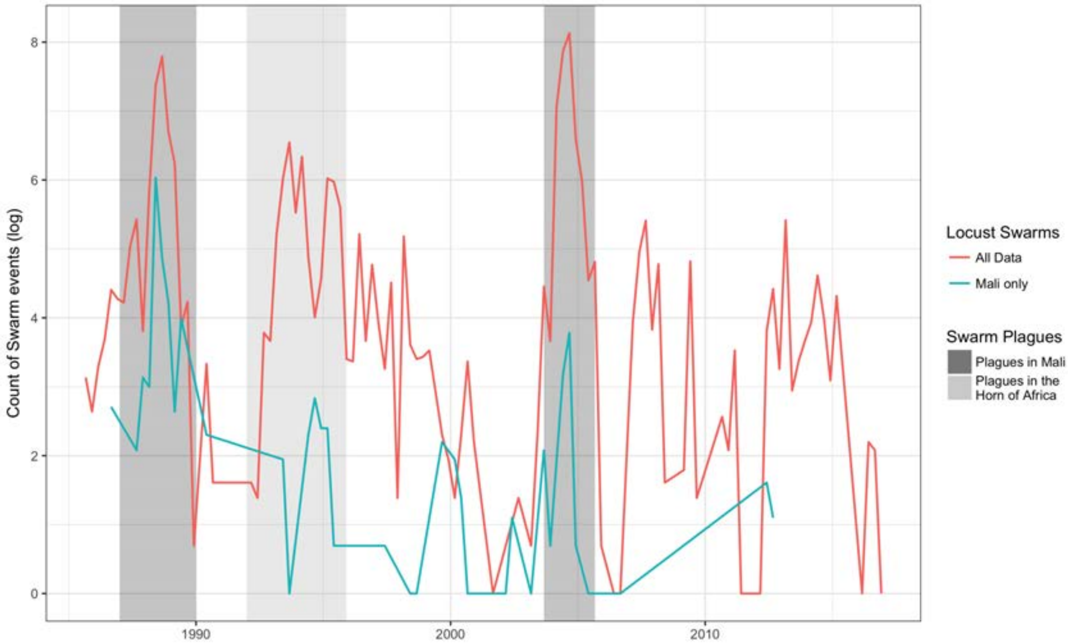
Appendix: Chapter 2

Figure B1: Geographical Distribution of Locust Swarms from 1985 to 2016.



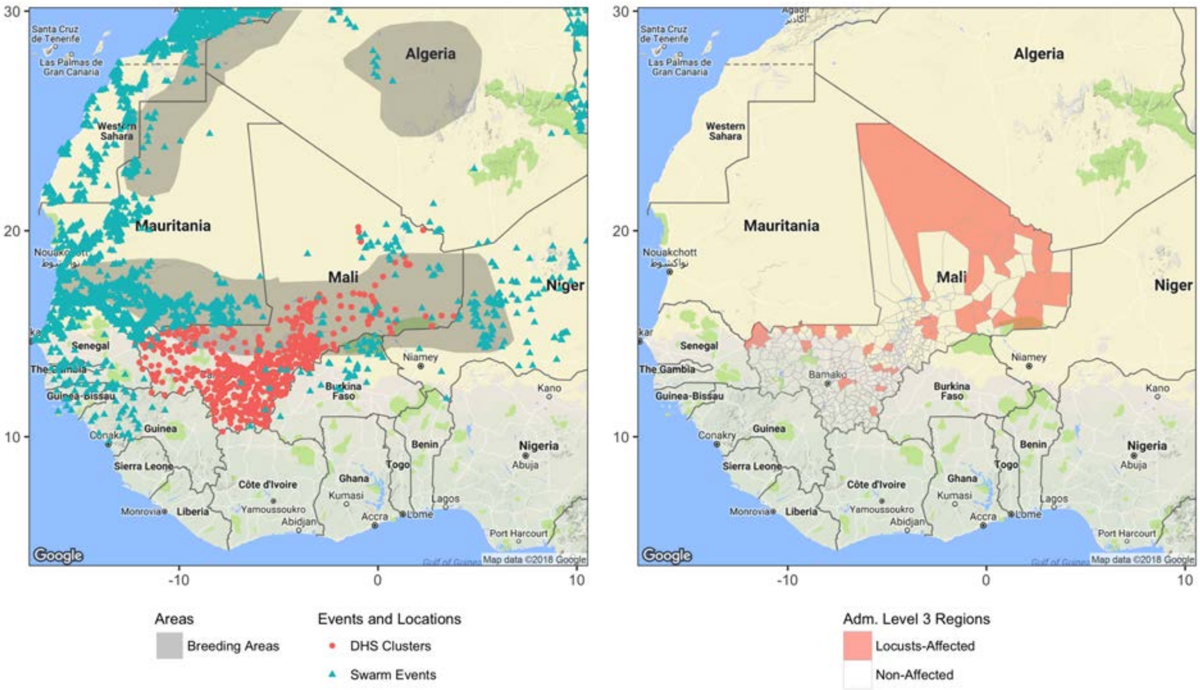
Source: SWARMS data base from FAO Desert Locust Information Service. See Section 2.3.

Figure B2: Quarterly Sum of Locust Swarms (in logs) for all regions and Mali only from 1985 to 2016.



Source: SWARMS data base from FAO Desert Locust Information Service. See Section 2.3.

Figure B3: Locust swarms' plague of 2003 to 2005 in Mali and the locations of DHS Clusters (left) and plague non-/affected areas (right).



Source: SWARMS data base from FAO Desert Locust Information Service (swarm events) and DHS Surveys V and VI for Mali (DHS clusters). The left panel plots the location of locust swarms and DHS clusters. The right panel the plague non-/affected areas, defined as Adm3-Level Regions. An area is defined as affected if a swarm event dated from 2003 to 2005 was spotted within its boundaries.

Appendix: Chapter 3

C.1 Robustness Table

Table C1: Executive Power and Ethnic Favoritism using log light as dependent variable

VARIABLES	(1) lnlight	(2) lnlight	(3) lnlight	(4) lnlight	(5) lnlight
leader	0.064*	0.090***			
	(0.037)	(0.021)			
leader * strong executive constraints		-0.068***			
		(0.015)			
senior partner			0.050***	0.064***	0.065***
			(0.018)	(0.024)	(0.024)
dominant			0.077**	0.092**	0.089**
			(0.031)	(0.042)	(0.042)
monopoly			0.153***	0.207***	0.211***
			(0.039)	(0.044)	(0.045)
senior partner * strong executive constraints				-0.043**	-0.041**
				(0.020)	(0.020)
dominant * strong executive constraints				-0.059*	-0.063**
				(0.031)	(0.030)
monopoly* strong executive constraints				-0.191***	-0.200***
				(0.051)	(0.052)
log population	-0.041	-0.038	-0.052	-0.023	0.018
	(0.296)	(0.138)	(0.140)	(0.140)	(0.140)
Observations	9,210	9,210	9,210	9,210	9,177
R-squared	0.996	0.996	0.996	0.996	0.996
Country-Year FE	YES	YES	YES	YES	YES
Group FE	YES	YES	YES	YES	YES
(population, area and urbanisation) x trend	NO	NO	NO	NO	YES

Robust standard errors clustered at country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. "lnlight" is the ln of night light emissions. "senior partner", "dominant" and "monopoly" are dummies that capture increasing executive power of the ethnic group. "leader" is a dummy that takes 1 if ethnic group has the highest level of access to power in the country/year ("senior partner", "dominant" or "monopoly"). "strong executive constraints" is a dummy which captures values of xconst>5. Regressions use population weights.

C.2 Data

C.2.1 Geographical Research On War, Unified Platform ($Grow^{up}$) Dataset

$Grow^{up}$ federated data platform provides access to disaggregated, integrated and spatially explicit conflict related data. It offers research-ready data on ethnic groups and intrastate conflict compiled from various sources and provided in group-year and country-year format.

The sample universe of ethnic groups in the RFE group-level data is adopted from the EPR (Ethnic Power Relations) Core dataset (Cederman et al., 2010). It covers all countries between 1946 and 2013 except failed states, overseas colonies and countries with less than 500'000 inhabitants. Newly independent states are included in the dataset beginning with the year of independence. From this country-year list, EPR Core dataset defines ethnicity as "any subjectively experienced sense of commonality based on the belief in common ancestry and shared culture". Only politically relevant ethnic groups are included in the dataset. An ethnic group is classified as relevant if "at least one political organization claims to represent it in national politics or if its members are subjected to state-led political discrimination". This yields 817 politically relevant ethnic groups in 141 countries in the EPR Core dataset.

Access to Power of Ethnic Groups

The data provided by Girardin et al. (2015) codes access to power on a seven point scale. Access to power is measured through the inclusion of members from the ethnic group in government. The different codes are:

- Monopoly: Elite members hold monopoly power in the executive to the exclusion of members of other ethnic groups.

- Dominant: Elite members of the group hold dominant power in the executive but there is limited inclusion of "token" members of other groups.

- Senior Partner: Representatives of the group participate as senior partners in a formal or informal power-sharing arrangement.
- Junior Partner: Representatives participate as junior partners in government.
- Self-Exclusion: The special category of self-exclusion applies to groups that have excluded themselves from central state power, in the sense that they control a particular territory of the state which they have declared independent from the central government.
- Powerless: Elite representatives hold no political power at either the national or the regional level without being explicitly discriminated against.
- Discriminated: Group members are subjected to active, intentional, and targeted discrimination, with the intent of excluding them from both regional and national power.

Night Light Data

Geographical data (night light intensity and population) at ethnic group level are raster derived data, created by overlaying the GeoEPR 2014 settlement polygons with geospatial raster datasets. Night light intensity data is taken from DMSP-OLS [1] Nighttime Lights Time Series (Average Visible, Stable Lights, and Cloud Free Coverages). All nightlights within group polygon are aggregated.

C.2.2 Polity IV Dataset

The Polity datasets aim at coding authority characteristics of states in the world system for purposes of quantitative analysis. The Polity IV dataset covers all major (around 167 countries), independent states in the global system over the period 1800-2014.

[1] Defense Meteorological Satellite Program

Bibliography

Acemoglu, Daron and James A Robinson, *Economic origins of dictatorship and democracy*, Cambridge University Press, 2005.

Akresh, Richard, German Caruso, and Harsha Thirumurthy, "Detailed Geographic Information, Conflict Exposure, and Health Impacts," 2016.

_ **, Leonardo Lucchetti, and Harsha Thirumurthy**, "Wars and Child Health: Evidence from the Eritrean–Ethiopian Conflict," *Journal of Development Economics*, 2012, *99* (2), 330–340.

_ **, Philip Verwimp, and Tom Bundervoet**, "Civil war, crop failure, and child stunting in Rwanda," *Economic Development and Cultural Change*, 2011, *59* (4), 777–810.

_ **, Sonia Bhalotra, Marinella Leone, and Una Okonkwo Osili**, "War and Stature: Growing Up During the Nigerian Civil War," *The American Economic Review (Papers and Proceedings)*, 2012, *102* (3), 273–277.

Alderman, Harold, John Hoddinott, and Bill Kinsey, "Long term consequences of early childhood malnutrition," *Oxford economic papers*, 2006, *58* (3), 450–474.

Alesina, Alberto and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 2005, *43* (3), 762–800.

Alesina, Alberto F, Stelios Michalopoulos, and Elias Papaioannou, "Ethnic inequality," *Journal of Political Economy, Forthcoming*, 2015.

Alesina, Alberto, Reza Baqir, and William Easterly, "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 1999, *114* (4).

Almond, Douglas and Janet Currie, "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives*, September 2011, *25* (3), 153–72.

_ , **Lena Edlund, Hongbin Li, and Junsen Zhang**, "Long-term effects of the 1959-1961 China famine: Mainland China and Hong Kong," Technical Report, National Bureau of Economic Research 2007.

Arias, María Alejandra, Ana Ibáñez, and Andres Zambrano, "Agricultural Production Amid Conflict: The Effects of Shocks, Uncertainty, and Governance of Non-State Armed Actors," Technical Report, Universidad de Los Aandes-Cede 2014.

Baker, Richard B, John Blanchette, and Katherine Eriksson, "Long-run Impacts of Agricultural Shocks on Educational Attainment: Evidence from the Boll Weevil," Technical Report, National Bureau of Economic Research 2018.

Banerjee, Abhijit, Esther Duflo, Gilles Postel-Vinay, and Tim Watts, "Long Run Health Impacts of Income Shocks: Wine and Phylloxera in 19th Century France," *The Review of Economics and Statistics*, 2010, *92* (4), 714–728.

Behrman, Jere R and Mark R Rosenzweig, "Returns to Birthweight," *Review of Economics and Statistics*, 2004, *86* (2), 586–601.

Berman, Nicolas, Mathieu Couttenier, and Dominic Rohner, "This Mine is Mine! How Minerals Fuel Conflicts in Africa," *American Economic Review*, 2017, *107* (6), 1564–1610.

Besley, Timothy and Hannes Mueller, "Estimating the Peace Dividend: The Impact of Violence on House Prices in Northern Ireland," *The American Economic Review*, 2012, *102* (2), 810–833.

_ **and Torsten Persson**, "The Logic of Political Violence," *The Quarterly Journal of Economics*, 2011, *126* (3), 1411–1445.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes, "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes," *The Quarterly Journal of Economics*, 2007, *122* (1), 409–439.

Blattman, Christopher, Alexandra Hartman, and Robert Blair, "How to promote order and property rights under weak rule of law? An experiment in changing dispute resolution behavior through community education," *American Political Science Review*, 2014, *108* (1), 100–120.

Bozzoli, Carlos and Tilman Brück, "Agriculture, Poverty, and Postwar Reconstruction: Micro-Level Evidence from Northern Mozambique," *Journal of Peace Research*, 2009, *46* (3), 377–397.

Buhaug, Halvard, Lars-Erik Cederman, and Kristian Skrede Gleditsch, "Square Pegs in Round Holes: Inequalities, Grievances, and Civil War," *International Studies Quarterly*, 2014, *58* (2), 418–431.

Bundervoet, Tom, Philip Verwimp, and Richard Akresh, "Health and Civil War in Rural Burundi," *Journal of Human Resources*, 2009, *44* (2), 536–563.

Buonanno, Paolo, Daniel Montolio, and Josep Maria Raya-Vílchez, "Housing Prices and Crime Perception," *Empirical Economics*, 2013, *45* (1), 305–321.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padro i Miquel, "The Value of Democracy: Evidence from Road Building in Kenya," *American Economic Review*, 2015, *105* (6), 1817–51.

Burke, Marshall, Solomon M Hsiang, and Edward Miguel, "Climate and conflict," *Annual Review of Economics*, 2015, *7* (1), 577–617.

Camacho, Adriana, "Stress and Birth Weight: Evidence from Terrorist Attacks," *American Economic Review: Papers and Proceedings*, 2008, *98* (2), 511–15.

Caselli, Francesco and Wilbur John Coleman, "On the theory of ethnic conflict," *Journal of the European Economic Association*, 2013, *11* (s1), 161–192.

Ceccato, Pietro, Keith Cressman, Alessandra Giannini, and Sylwia Trzaska, "The desert locust upsurge in West Africa (2003–2005): Information on the desert locust early warning system and the prospects for seasonal climate forecasting," *International Journal of Pest Management*, 2007, *53* (1), 7–13.

Cederman, Lars-Erik, Andreas Wimmer, and Brian Min, "Why Do Ethnic Groups Rebel? New Data and Analysis," *World Politics*, 2010, *62*, 87–119.

_ **, Kristian Skrede Gleditsch, and Halvard Buhaug**, *Inequality, grievances, and civil war*, Cambridge University Press, 2013.

Chainey, Spencer, Lisa Tompson, and Sebastian Uhlig, "The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime," *Security Journal*, 2008, *21* (1-2), 4–28.

Collier, Paul and A Hoeffler, "Greed and grievance: Economic agendas in civil wars," *Lynne Rienner, chapter Doing Well Out of War: An Economic Perspective*, 2000, p. 91.

Cressman, K, "SWARMS: a geographic information system for desert locust forecasting," in "New Strategies in locust control," Springer, 1997, pp. 27–35.

Cressman, Keith and Robert Stefanski, "Weather and Desert Locusts," Technical Report, World Meteorological Organization and Food and Agriculture Organization of the United Nations 2016.

Dagnelie, Olivier, Giacomo De Luca, and Jean-François Maystadt, "Violence, Selection and Infant Mortality in Congo," *Journal of Health Economics*, 2018.

Deininger, Klaus, "Causes and Consequences of Civil Strife: Micro-Level Evidence from Uganda," *Oxford Economic Papers*, 2003, pp. 579–606.

DiNardo, John, Nicole M Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, September 1996, *64* (5), 1001–1044.

Easterly, William and Ross Levine, "Africa's growth tragedy: policies and ethnic divisions," *The Quarterly Journal of Economics*, 1997, pp. 1203–1250.

Esteban, Joan and Debraj Ray, "On the Salience of Ethnic Conflict," *American Economic Review*, 2008, *98* (5), 2185–2202.

__ , **Laura Mayoral, and Debraj Ray**, “Ethnicity and Conflict: An Empirical Study,” *American Economic Review*, 2012, *102* (4), 1310–42.

FAO, “Locust crisis to hit northwest Africa again - situation deteriorating in the Sahel (last visit 05.06.2018),” September 2004.

__ , “West African food crisis looming (last visit 05.06.2018),” June 2005.

__ , “Worsening food situation in parts of the Sahel (last visit 05.06.2018),” May 2005.

Fearon, James D. and David D. Laitin, “Violence and the Social Construction of Ethnic Identity,” *International Organization*, 9 2000, *54*, 845–877.

Fürst, Thomas, Giovanna Raso, Cinthia A Acka, Andres B Tschannen, Eliézer K N’Goran, and Jürg Utzinger, “Dynamics of Socioeconomic Risk Factors for Neglected Tropical Diseases and Malaria in an Armed Conflict,” *PLOS Neglected Tropical Diseases*, 2009, *3* (9), e513.

Gerber, Matthew S, “Predicting Crime using Twitter and Kernel Density Estimation,” *Decision Support Systems*, 2014, *61*, 115–125.

Gibbons, Steve, “The Costs of Urban Property Crime,” *The Economic Journal*, 2004, *114* (499).

Girardin, Luc, Hunziker Philipp, Cederman Lars-Erik, Bormann Nils-Christian, and Vog Manuel, 2015. *GROW^{up}* - Geographical Research On War, Unified Platform. ETH Zurich. http://growup.ethz.ch/.

Goldstone, Jack A, Robert H Bates, David L Epstein, Ted Robert Gurr, Michael B Lustik, Monty G Marshall, Jay Ulfelder, and Mark Woodward, “A global model for forecasting political instability,” *American Journal of Political Science*, 2010, *54* (1), 190–208.

Harari, Mariaflavia and Eliana La Ferrara, “Conflict, Climate and Cells: A Disaggregated Analysis,” Technical Report, Working Paper 2017.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, "Unsupervised Learning," in "The Elements of Statistical Learning," Springer, 2009, pp. 485–585.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil, "Measuring Economic Growth from Outer Space," *American Economic Review*, 2012, *102* (2), 994–1028.

Herok, Claudia A, Stephan Krall et al., *Economics of Desert locust control.*, Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ) GmbH, 1995.

Hodler, Roland and Paul A. Raschky, "Regional Favoritism," *The Quarterly Journal of Economics*, 2014, *129* (2), 995–1033.

Hsiang, Solomon M, Marshall Burke, and Edward Miguel, "Quantifying the Influence of Climate on Human Conflict," *Science*, 2013, *341* (6151), 1235367.

Hu, Yujie, Fahui Wang, Cecile Guin, and Haojie Zhu, "A Spatio-Temporal Kernel Density Estimation Framework for Predictive Crime Hotspot Mapping and Evaluation," *Applied Geography*, 2018, *99*, 89–97.

Izenman, Alan Julian, *Modern Multivariate Statistical Techniques: Regression, Classification and Manifold learning*, Springer, 2008.

Kudamatsu, Masayuki, "Ethnic favoritism: Micro evidence from Guinea," 2009.

_ , "Has Democratization Reduced Infant Mortality in Sub-Saharan Africa? Evidence from Micro Data," *Journal of the European Economic Association*, 2012, *10* (6), 1294–1317.

_ , Torsten Persson, and David Strömberg, "Weather and Infant Mortality in Africa," Technical Report, CEPR Discussion Papers 9222 2012.

Lavy, Victor, Analia Schlosser, and Adi Shany, "Out of Africa: Human Capital Consequences of In Utero Conditions," Working Paper 21894, National Bureau of Economic Research January 2016.

Leon, Gianmarco, "Civil Conflict and Human Capital Accumulation: The Long Term Effects of Political Violence in Perú," *Journal of Human Resources*, 2012, *47* (4), 991–1022.

Luca, Giacomo De, Roland Hodler, Paul A Raschky, and Michele Valsecchi, "Ethnic Favoritism: An Axiom of Politics?," *CESifo Working Paper Series*, 2015.

Maccini, Sharon and Dean Yang, "Under the Weather: Health, Schooling, and Economic Consequences of Early Life Rainfall," *American Economic Review*, 2009, *99* (3), 1006–26.

Maluccio, John A, John Hoddinott, Jere R Behrman, Reynaldo Martorell, Agnes R Quisumbing, and Aryeh D Stein, "The Impact of Improving Nutrition During Early Childhood on Education among Guatemalan Adults," *The Economic Journal*, 2009, *119* (537), 734–763.

Mansour, Hani and Daniel I Rees, "Armed Conflict and Birth Weight: Evidence from the Al-Aqsa Intifada," *Journal of Development Economics*, 2012, *99* (1), 190–199.

Marshall, Monty G. and Ted Robert Gurr, 2013. Polity IV project - Center for International Development and Conflict Management at the University of Maryland College Park. http://www.systemicpeace.org/polity/polity4.htm.

Matsuura, K and CJ Willmott, "Terrestrial Air Temperature and Precipitation: 1900–2014 Gridded Time Series (version 4.01)," 2015.

Michalopoulos, Stelios and Elias Papaioannou, "Pre-Colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, 2013, *81* (1), 113–152.

_ **and** _ , "The long-run effects of the scramble for Africa," *American Economic Review, Forthcoming*, 2016.

Mill, John Stuart, *On liberty*, Longmans, Green, Reader, and Dyer, 1869.

Minoiu, Camelia and Olga N Shemyakina, "Armed conflict, Household Victimization, and Child Health in Côte d'Ivoire," *Journal of Development Economics*, 2014, *108*, 237–255.

— **and Sanjay G Reddy**, "Kernel Density Estimation on Grouped data: The Case of Poverty Assessment," *The Journal of Economic Inequality*, 2014, *12* (2), 163–189.

Montalvo, Jose G. and Marta Reynal-Querol, "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review*, 2005, *95* (3), 796–816.

Mueller, Hannes Felix, Lavinia Piemontese, and Augustin Tapsoba, "Recovery from Conflict: Lessons of Success," *World Bank Policy Research Paper*, 2017, (7970).

NOAA-NGDC, 2013. Version 4 DMSP-OLS Nighttime Lights Time Series (Average Visible, Stable Lights, & Cloud Free Coverages). National Oceanic and Atmospheric Administration-National Geophysical Data Center. http://ngdc.noaa.gov/eog/dmsp.html.

Ouili, Idrissa, "Armed Conflicts, Children's Education and Mortality: New Evidence from Ivory Coast," *Journal of Family and Economic Issues*, 2017, *38* (2), 163–183.

Pongou, Roland, "Why Is Infant Mortality Higher in Boys than in Girls? A new Hypothesis Based on Preconception Environment and Evidence from a Large Sample of Twins," *Demography*, 2013, *50* (2), 421–444.

Pope, Jaren C, "Fear of Crime and Housing Prices: Household Reactions to Sex Offender Registries," *Journal of Urban Economics*, 2008, *64* (3), 601–614.

Rockmore, Marc, "The Cost of Fear: The Welfare Effect of the Risk of Violence in Northern Uganda," *The World Bank Economic Review*, 2017, *31* (3), 650–669.

Rodrik, Dani, "Where did all the growth go? External shocks, social conflict, and growth collapses," *Journal of economic growth*, 1999, *4* (4), 385–412.

Sala-i-Martin, Xavier, "The World Distribution of Income (Estimated from Individual Country Distributions)," Technical Report, National Bureau of Economic Research 2002.

— , "The World Distribution of Income: Falling Poverty and ...Convergence, Period," *The Quarterly Journal of Economics*, 2006, *121* (2), 351–397.

Sany, Joseph, *Education and Conflict in Côte D'Ivoire*, United States Institute of Peace, 2010.

Shemyakina, Olga, "Exploring the Impact of Conflict Exposure during Formative Years on Labour Market Outcomes in Tajikistan," *The Journal of Development Studies*, 2015, *51* (4), 422–446.

Silverman, Bernard W, *Density Estimation for Statistics and Data Analysis*, Vol. 26, CRC press, 1986.

Stein, Zena, Mervyn Susser, Gerhart Saenger, and Francis Marolla, *Famine and Human Development: The Dutch Hunger Winter of 1944-1945.*, Oxford University Press, 1975.

Symmons, PM, K Cressman, and HM Dobson, "Desert locust guidelines," 2001.

Taylor, Charles C, "Bootstrap Choice of the Smoothing Parameter in Kernel Density Estimation," *Biometrika*, 1989, *76* (4), 705–712.

Terrell, George R and David W Scott, "Variable Kernel Density Estimation," *The Annals of Statistics*, 1992, pp. 1236–1265.

Tollefsen, Andreas Forø, Håvard Strand, and Halvard Buhaug, "PRIO-GRID: A Unified Spatial Sata Structure," *Journal of Peace Research*, 2012, *49* (2), 363–374.

UNOCHA, "Fighting in Côte d'Ivoire jeopardizes humanitarian aid," https://www.un.org/press/en/2004/afr1061.doc.htm 2004.

Valente, Christine, "Civil Conflict, Gender-specific Fetal Loss, and Selection: A new Test of the Trivers–Willard hypothesis," *Journal of Health Economics*, 2015, *39*, 31–50.

Vreyer, Philippe De, Nathalie Guilbert, and Sandrine Mesple-Somps, "Impact of natural disasters on education outcomes: evidence from the 1987–89 locust plague in Mali," *Journal of African Economies*, 2014, *24* (1), 57–100.

Walque, Damien De, "Selective Mortality during the Khmer Rouge Period in Cambodia," *Population and Development Review*, 2005, *31* (2), 351–368.

Willard, Kristen L, Timothy W Guinnane, and Harvey S Rosen, "Turning Points in the Civil War: Views from the Greenback Market," *American Economic Review*, September 1996, *86* (4), 1001–1018.

Zussman, Asaf, Noam Zussman, and Morten Ørregaard Nielsen, "Asset Market Perspectives on the Israeli–Palestinian Conflict," *Economica*, 2008, *75* (297), 84–115.